



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 6:</b> C12Q 1/68, C12P 19/34, C07H 21/02, 21/04	<b>A1</b>	<b>(11) International Publication Number:</b> WO 97/45560 <b>(43) International Publication Date:</b> 4 December 1997 (04.12.97)
<b>(21) International Application Number:</b> PCT/US97/09599 <b>(22) International Filing Date:</b> 30 May 1997 (30.05.97)  <b>(30) Priority Data:</b> 08/657,866 31 May 1996 (31.05.96) US  <b>(71) Applicant:</b> NORTH SHORE UNIVERSITY HOSPITAL RESEARCH CORPORATION [US/US]; 380 Community Drive, Manhasset, NY 11030 (US).  <b>(72) Inventor:</b> DUFFY, Hao-peng, Xu; 29 North Drive, Centerport, NY 11721 (US).  <b>(74) Agents:</b> SINDER, Stuart, J. et al.; Kenyon & Kenyon, 1 Broadway, New York, NY 10004 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> IDENTIFICATION OF DIFFERENTIALLY METHYLATED AND MUTATED NUCLEIC ACIDS  <b>(57) Abstract</b>  The present invention provides Methyl- (or Mutant-) Differential Display (MDD) methods and nucleic acid probes for detecting mutations and the methylation patterns of nucleic acids. The methods of the present invention are particularly for detecting and isolating genomic DNA fragments which are near coding and regulatory regions of genes and which are differentially mutated or methylated relative to the corresponding DNA from normal cells. Genes are frequently not methylated in the cells where they are expressed but are methylated in cell types where they are not expressed. Moreover, tumor cell DNA is frequently methylated to a different extent and in different regions than is the DNA of normal cells. The present invention is used for identifying which regions of the genome are methylated or mutated in different cell types, including cancerous cell types. The present invention is also used for diagnosing whether a tissue sample is cancerous, and whether that cancerous condition is non-metastatic or metastatic.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

IDENTIFICATION OF DIFFERENTIALLY METHYLATED  
AND MUTATED NUCLEIC ACIDS

5

FIELD OF THE INVENTION

The present invention provides Methyl- (or Mutant-) Differential Display (MDD) methods and nucleic acid probes for detecting mutations and/or the methylation patterns of nucleic acids. Genes are frequently not methylated in the cells where they are expressed but are methylated in cell types where they are not expressed. Moreover, tumor cell DNA is frequently methylated to a different extent and in different regions than is the DNA of normal cells. The present invention is used for identifying which regions of the genome are methylated or mutated in different cell types, including cancerous cell types. The present invention also is used for identifying whether a tissue sample has a methylation or mutation pattern which is normal or like that of known cancer cells.

BACKGROUND OF THE INVENTION

DNA is often methylated in normal mammalian cells. For example, DNA is methylated to determine whether a given gene will be expressed and whether the maternal or the paternal allele of that gene will be expressed. See Melissa Little et al., *Methylation and p16: Suppressing the Suppressor*, 1 NATURE MEDICINE 633 (1995). While methylation is known to occur at CpG sequences, only recent studies indicate that CpNpG sequences may be methylated. Susan J. Clark et al., *CpNpGp Methylation in Mammalian Cells*, 10 NATURE GENETICS 20, 20 (1995). Methylation at CpG sites has been much more widely studied and is better understood.

Methylation occurs by enzymatic recognition of CpG and CpNpG sequences followed by placement of a methyl (CH<sub>3</sub>) group on the fifth carbon atom of a cytosine base. The enzyme that mediates methylation of CpG dinucleotides, 5-cytosine methyltransferase, is essential for embryonic development -- without it embryos die soon after gastrulation. It is not yet clear whether this enzyme methylates CpNpG sites. Peter W. Laird et al., *DNA Methylation and Cancer*, 3 HUMAN MOLECULAR GENETICS 1487, 1488 (1994).

When a gene has many methylated cytosines it is less likely to be expressed. K Willson, 7 TRENDS GENET. 107-109 (1991). Hence, if a maternally-inherited gene is more highly methylated than the paternally-inherited gene, the paternally-inherited gene will give rise to more gene product. Similarly, when a gene is expressed in a tissue-specific manner, that gene will often be unmethylated in the tissues where it is active, but will be highly methylated in the tissues where it is inactive. Incorrect methylation is thought to be the cause of some diseases including Beckwith-Wiedemann syndrome and Prader-Willi syndrome. I. Henry et al., 351 NATURE 665, 667 (1991); R.D. Nicholls et al., 342 NATURE 281, 281-85 (1989).

The methylation patterns of DNA from tumor cells are generally different than those of normal cells. Laird et al., *supra*. Tumor cell DNA is generally undermethylated relative to normal cell DNA, but selected regions of the tumor cell genome may be more highly methylated than the same regions of a normal cell's genome. Hence, detection of altered methylation patterns in the DNA of a tissue sample is an indication that the tissue is cancerous. For example, the gene for Insulin-Like Growth Factor 2 (IGF2) is hypomethylated in a number of cancerous tissues, such as Wilm's Tumors, rhabdomyosarcoma, lung cancer and hepatoblastomas. Rainner et al. 362 NATURE 747-49 (1993); Ogawa, et al., 362 NATURE 749-51 (1993); S. Zhan et al., 94 J. CLIN. INVEST. 445-48 (1994); P.V. Pedone et al., 3 HUM. MOL. GENET. 1117-21 (1994); H. Suzuki et al., 7 NATURE GENET 432-38 (1994); S. Rainier et al., 55 CANCER RES 1836-38 (1995).

The present invention is directed to a method of detecting differential methylation at CpNpG sequences by cutting test and control DNAs with a restriction enzyme that will not cut methylated DNA, and then detecting the difference in size of the resulting restriction fragments.

While methylation-sensitive restriction enzymes have been used for observing differential methylation in various cells, no commercial assays exist for use on human samples because differentially methylated sequences represent such a minute proportion of the human genome that they are not readily detected. The human genome is both highly complex, in that it contains a great diversity of DNA sequences, and highly repetitive, in that it contains a lot of DNA with very similar or identical sequences. The high complexity and repetitiveness of human DNA confounds efforts at detecting and isolating the minute amount of differentially methylated DNA which may be present in a

test sample. The present invention remedies the detection problem by providing new procedures for screening a selected subset of the mammalian genome which is most likely to contain genetic functions.

5 The present invention provides techniques for detecting and isolating differentially methylated or mutated segments of DNA which may be present in a tissue sample in only minute amounts by using one or more rounds of DNA amplification coupled with subtractive hybridization to identify such segments of DNA. DNA amplification has been coupled with subtractive hybridization in the Representational Difference Analysis (RDA) procedures disclosed in U.S. Patent 5,436,142 to Wigler et al. and Nikolai Lisitsyn et al.,  
10 *Cloning the Differences Between Two Complex Genomes*, 259 SCIENCE 946 (1993). However, for the subtractive hybridization step of such RDA procedures to proceed in a reasonable time and with reasonable efficiency, only a subset of the genome can be examined. To accomplish this necessary reduction in the complexity of the sample DNA, Wigler et al. and Lisitsyn et al. disclose cutting DNA samples with restriction enzymes that  
15 cut infrequently and randomly. However, selection of enzymes which randomly cut the genome, means that the portion of the genome which is examined is not enriched for any particular population of DNA fragments. Thus, when RDA is used, only a random subset of the human genome, which includes repetitive elements, noncoding regions and other sequences which are generally not of interest, can be tested in a single experiment.

20 In contrast, the present invention is directed to methods which use enzymes that cut frequently and that specifically cut CG-rich regions of the genome. These enzymes are chosen because CG-rich regions of the genome are not evenly distributed in the genome -- instead, CG-rich regions are frequently found near genes, and particularly near the promoter regions of genes. This means that the proportion of the genome that is  
25 examined by the present methods will be enriched for genetically-encoded sequences as well as regulatory sequences. Moreover, unlike the RDA method, the present methods selectively identify regions of the genome which are hypomethylated or hypermethylated by using enzymes which specifically cut non-methylated CG-rich sequences. The present invention therefore represents an improvement over RDA methods because of its ability  
30 to select DNA fragments which are likely to be near or to encode genetic functions.

#### SUMMARY OF THE INVENTION:

The present invention provides probes and methods of detecting whether a CNGtriplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cell.

5 The present invention provides a method of detecting whether a CNGtriplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cells which includes:

a) isolating genomic DNA from a control sample of cells and a test sample of cells to generate a control-cell DNA and a test-cell DNA;

10 b) cleaving the control-cell DNA and the test-cell DNA with a master restriction enzyme to generate cleaved control-cell DNA and cleaved test-cell DNA;

c) preparing a probe from a DNA isolated by the present methods, for example, a DNA selected from the group consisting of SEQ ID NO:7-10;

d) hybridizing the probe to the cleaved control-cell DNA and the cleaved test-cell DNA to form a control-hybridization complex and a test-hybridization complex; and

15 e) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex;

wherein the master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence

20 Similarly, the present invention provides a method of detecting whether a DNA site is mutated in a genomic DNA present in a test sample of cells which includes steps

a) through e) above but where a detector restriction enzyme is used instead of the master restriction enzyme,

wherein the detector restriction enzyme does not cleave a mutated DNA site but does cleave a corresponding nonmutated DNA site.

25 The present invention also provides methods of isolating probes to detect hypomethylation or hypermethylation in a CNGtriplet of DNA. In particular, the present invention provides a method of isolating a probe to detect hypomethylation in a CNGtriplet of DNA which includes:

30 a) cleaving a tester sample of genomic DNA with both a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

b) ligating a first set of adaptors onto master enzyme cut DNA ends of the cleaved tester sample to generate a first-tester amplification template;

c) amplifying the first-tester amplification template to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to the first set of adaptors;

5 d) cleaving off the first adaptors from the first-tester amplicon and ligating a second set of adaptors onto DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends;

e) melting and hybridizing the second-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a mixture of tester-tester product and tester-driver product;

10 f) adding nucleotides onto DNA ends present in the mixture to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

g) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-tester amplicon;

15 h) isolating a discrete DNA fragment from the second tester amplicon as a probe to detect the hypomethylation or the hypermethylation in a CNGtriplet of DNA;

wherein the master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence;

20 wherein the partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein the driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme and amplified using primers that recognize DNA ends cut by the master restriction enzyme.

25 Similarly, when the present invention provides a method of isolating a probe to detect hypermethylation in a CNGtriplet of DNA, normal DNA is used instead of the tester DNA and the normal amplicon is annealed and hybridized with a molar excess of tester driver DNA.

30 Moreover the present methods are used to isolate probes to detect a mutation in a test sample of genomic DNA by using a detector enzyme instead of a partner restriction enzyme, wherein the detector restriction enzyme cleaves a normal DNA site but not a mutant DNA site.

The present invention further provides a kit for detecting hypomethylation in a CNGtriplet of DNA which is present in a test tissue sample which includes a DNA or a probe isolated by the methods of the present invention, for example, a DNA having SEQ ID NO:7-10.

5

#### BRIEF DESCRIPTIONS OF THE DRAWINGS:

Fig. 1 provides a schematic diagram of how the MDD technique generally selects for DNA fragments which have non-methylated CG-rich ends. Driver DNA is isolated from a normal cell or tissue sample of the same tissue-type as the tester cell DNA. Tester DNA is isolated from a cell or tissue sample which likely has a mutation or a difference in the extent to which its DNA is methylated. Tester and driver DNAs are cleaved by the master (M) and partner (P) restriction enzymes. The master enzyme will not cleave DNA at sites that have a methyl group (superscript "m") but will cleave at sites which are non-methylated (superscript m having a line drawn through it). The partner enzyme cleaves the DNA into a size and a complexity appropriate for DNA amplification and subtractive hybridization. After cleavage, a small target fragment is completely excised from the tester DNA but not from the driver DNA. Adaptors (m-linkers) which hybridize only to master-enzyme-cut DNA ends are ligated onto the tester and driver DNA fragments. Unique target DNAs from the tester sample will have adaptors on both ends. Upon DNA amplification using a primer that recognizes only the adaptor sequences, only DNA fragments with adaptors on both ends will be efficiently amplified.

Fig. 2 depicts an agarose gel containing electrophoretically separated DNA fragments isolated by MDD from malignant B-cells derived from Chronic Lymphocytic Leukemia (CLL) patients. Lane A contains the different products isolated from CLL patient #111. Lane M contains Hae III 0174 DNA size markers. As illustrated, the present MDD methods give rise to discrete fragments of DNA which can readily be isolated and subcloned or used as probes for identifying the methylation patterns of genomic DNA samples from different cell types

Fig. 3 depicts a Southern Blot of MspI-digested tester (T) and driver (D) DNA hybridized with the CLL58 probe. The lane labeled "probe #58" contains a clone of the CLL58 probe. The lane labeled "T" contains amplified 2 $\mu$ g CD5<sup>+</sup> B-cell DNA, whereas the lane "D" contains 2 $\mu$ g amplified neutrophil DNA. Both T and D DNA samples were



isolated from the CLL patient #111 and amplified by the MDD method. As a positive control, probe CLL58 hybridizes strongly to itself. Probe CLL58 also hybridizes strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA hypomethylation in the tester DNA.

5           Fig. 4 depicts a Southern Blot of MspI-digested DNA hybridized with the CLL58 probe. The lane labeled "probe #58" contains a clone of the CLL58 probe. The lane labeled "T" contains 6  $\mu$ g of CD5<sup>+</sup> B-cell DNA, whereas lane "N" contains 6  $\mu$ g of neutrophil DNA, each isolated from the CLL patient #111. As a positive control, probe CLL58 hybridizes strongly to itself. Probe CLL58 hybridizes with a lower band and  
10 several upper bands in B-cell DNA (lane T), while comparatively hybridizing much less with the lower band in neutrophil cell DNA (lane N). The lesser hybridization to the lower band in neutrophil DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the CLL58 probe provides a comparative methyl-differential display in B-cells and neutrophil cells.

15           Fig. 5 depicts an agarose gel after electrophoresis of the different final products isolated from breast patient #13 (lane A), patient #14 (lane B) and breast patient #4 (lane C) using the MDD technique. Lane MW contains Hae III 0174 DNA size markers.

          Fig. 6a depicts a Southern Blot of MspI-digested tester (T) and driver (D) DNA hybridized with the BR50 probe. The lane labeled "probe #50" contains a clone of the  
20 BR50 probe. The lane labeled "T" contains amplified breast tumor cell DNA, whereas the lane "D" contains amplified normal breast cell DNA. Both T and D DNA samples were isolated from the breast patient #14 and amplified by the MDD method. As a positive control, probe BR50 hybridizes strongly to itself. Probe BR50 also hybridizes strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA  
25 hypomethylation in the tester DNA.

          Fig. 6b depicts a Southern Blot of MspI-digested tester (T) and driver (D) DNA hybridized with the BR104 probe. The lane labeled "probe #104" contains a clone of the  
BR104 probe. The lane labeled "T" contains amplified breast tumor cell DNA, whereas the lane "D" contains amplified normal breast cell DNA. Both T and D DNA samples  
30 were isolated from the breast patient #13 and amplified by the MDD method. As a positive control, probe BR104 hybridizes strongly to itself. Probe BR104 also hybridizes

strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA hypomethylation in the tester DNA.

Fig. 6c depicts a Southern Blot of MspI-digested tester (T) and driver (D) DNA hybridized with the BR254 probe. The lane labeled "probe #254" contains a clone of the BR254 probe. The lane labeled "T" contains amplified breast tumor cell DNA, whereas the lane "D" contains amplified normal breast cell DNA. Both T and D DNA samples were isolated from the breast patient #4 and amplified by the MDD method. As a positive control, probe BR254 hybridizes strongly to itself. Probe BR254 also hybridizes strongly with the tester DNA but only weakly with the driver amplicon DNA -- this is due to DNA hypomethylation in the tester DNA.

Fig. 7a depicts a genomic Southern Blot of MSpI-digested DNA hybridized with a probe BR50 isolated from a breast cancer patient. The lane labeled "probe #50" contains a clone of the isolated BR50 probe. The lane labeled "T" contains 6  $\mu$ g of B-cell DNA, whereas lane "N" contains 6  $\mu$ g of normal breast DNA, each isolated from the breast cancer patient #14. As a positive control, probe BR50 hybridizes strongly to itself. Probe BR50 hybridizes strongly with a lower band and weakly with a medium band in breast tumor DNA (lane T). However, in normal breast cell DNA (lane N), probe BR50 hybridizes much less with the lower band and only weakly to an upper band. The lesser hybridization to the lower band in normal breast DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the BR50 probe provides a comparative methyl-differential display in tumor and normal breast cells.

Fig 7b depicts a genomic Southern Blot of MSpI-digested DNA hybridized with a probe BR104 isolated from a breast cancer patient. The lane labeled "probe #104" contains a clone of the isolated BR104 probe. The lane labeled "T" contains 6  $\mu$ g of tumor DNA, whereas lane "N" contains 6  $\mu$ g of normal breast cell DNA, each isolated from the breast cancer patient #13. As a positive control, probe BR104 hybridizes strongly to itself. Probe BR104 hybridizes strongly with a lower band and weakly with a medium band in tumor DNA (lane T). However, in normal breast cell DNA (lane N), probe BR104 hybridizes lightly with the lower band and moderately with middle and upper bands. The lesser hybridization to the lower band in normal breast DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the BR104

probe provides a comparative methyl-differential display in tumorous and normal breast cells.

Fig. 7c depicts a genomic Southern Blot of MspI-digested DNA hybridized with a probe BR254 isolated from a breast cancer patient. The lane labeled "probe #254" contains a clone of the isolated BR254 probe. The lane labeled "T" contains 6  $\mu$ g of B-cell DNA, whereas lane "N" contains 6  $\mu$ g of normal breast cell DNA, each isolated from the breast cancer patient #4. As a positive control, probe BR254 hybridizes strongly to itself. Probe BR254 hybridizes strongly with a lower band and with an upper band in tumor DNA (lane T). However, in normal breast cell DNA (lane N), probe BR254 hybridizes only with the upper band. The lack of hybridization to the lower band in normal breast DNA is due to DNA methylation which prevents excision of that fragment. Hence, the BR254 probe provides a comparative methyl-differential display in tumorous and normal breast cells.

Fig. 8 depicts a Southern blot of MspI-digested genomic DNA from five different breast cancer patients which was hybridized with the BR50 probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #50" contains a clone of the isolated BR50 probe. The lane labeled "T" contains 6  $\mu$ g of DNA isolated from a tumor, whereas lane "N" contains 6  $\mu$ g of DNA from normal breast tissue. As a positive control, probe BR50 hybridizes strongly to itself. For each of the five patients, probe BR50 hybridizes strongly with a lower band and less strongly with an upper band in the tumor DNA samples (T). However, probe BR50 hybridizes much less with the lower band in the normal DNA samples (N). The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the BR50 probe provides a comparative methyl-differential display in breast tumor and normal DNA samples.

Fig. 9 depicts a Southern blot of MspI-digested genomic DNA from four different breast cancer patients which was hybridized with the BR104 probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #104" contains a clone of the isolated BR104 probe. As a positive control, probe BR104 hybridizes strongly to itself. The lanes labeled "T" contain 6  $\mu$ g of DNA isolated

from a breast tumor, whereas the lanes labeled "N" contain 6  $\mu$ g of DNA from normal breast tissue. For tumor DNA from each of the four patients, probe BR104 hybridizes with a lower band and a middle band. In the normal DNA, the probe hybridizes mainly with the middle band and an upper band. The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the BR104 probe provides a comparative methyl-differential display in breast tumor and normal breast DNA samples.

Fig. 10 depicts a Southern blot of MspI-digested genomic DNA from four different breast cancer patients which was hybridized with the BR254 probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #254" contains a clone of the isolated BR254 probe. The lane labeled "T" contains 6  $\mu$ g of DNA isolated from a breast tumor, whereas lane "N" contains 6  $\mu$ g of DNA from normal breast tissue. As a positive control, probe BR254 hybridizes strongly to itself. For each of the four patients, probe BR254 hybridizes strongly with a lower band and less strongly with an upper band in the tumor DNA samples (T). However, probe BR254 hybridizes much less with the lower band in the normal DNA samples (N). The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the BR254 probe provides a comparative methyl-differential display in breast tumor and normal DNA samples.

Fig. 11 depicts an amplification pattern detected by probe BR254 in DNA isolated from a breast tumor biopsy. Fig. 12 provides a Southern blot of normal (N) and tumor (T) DNA from breast cancer patient #23 which was digested with MspI and hybridized with the BR254 probe. Both the N and T lanes contain 6  $\mu$ g of DNA. The lane labeled "probe #254" contains a clone of the isolated BR254 probe, which hybridizes to itself. In the tumor DNA, the BR254 probe hybridizes very, very strongly to a DNA fragment to which it hybridizes only moderately in normal DNA. This indicates that the breast tumor DNA of patient #23 has a highly amplified region of genomic DNA which is not present in her normal DNA.

Fig. 12 depicts a Southern blot of MspI-digested genomic DNA from five different ovarian cancer patients which was hybridized with the BR50 probe. The number for each patient is listed above the bracket. The hybridization patterns of 6  $\mu$ g primary tumor

DNA (T or T1), 6  $\mu$ g metastatic tumor DNA (Tm) and 6  $\mu$ g normal ovarian DNA (N) from these patients are provided for comparison. The lane labeled "probe #50" contains a clone of the isolated BR50 probe. The lane labeled "T" contains 6  $\mu$ g of DNA isolated from a ovarian tumor, whereas lane "N" contains 6  $\mu$ g of DNA from normal ovarian tissue. As a positive control, probe BR50 hybridizes strongly to itself. In primary tumor DNAs, the probe predominantly hybridizes with the lower bands, and slightly hybridizes with the upper bands. In metastatic tumor DNAs, the probe only hybridizes with the lower bands. In normal DNAs, the probe hybridized with the lower bands and the upper bands. The lesser hybridization to the lower band in normal DNA is due to DNA methylation which prevents complete excision of that fragment. Hence, the BR50 probe provides a comparative methyl-differential display in ovarian tumor and normal DNA samples. The degree of DNA hypomethylation may be related to the progress of the disease.

Fig. 13 depicts a Southern blot of MspI-digested genomic DNA from four different ovarian cancer patients which was hybridized with the BR254 probe. The number for each patient is listed above the bracket. The hybridization patterns of tumor (T) and matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #254" contains a clone of the isolated BR254 probe. The lane labeled "T" contains 6  $\mu$ g of DNA isolated from a ovarian tumor, whereas lane "N" contains 6  $\mu$ g of DNA from normal ovarian tissue. As a positive control, probe BR254 hybridizes strongly to itself. For each of the four patients, probe BR254 hybridizes with a lower band and with an upper band in the tumor DNA samples (T). However, probe BR254 hybridizes only with the upper band in the normal DNA samples (N). The lack of hybridization to the lower band in normal DNA is due to DNA methylation which prevents excision of that fragment. Hence, the BR254 probe provides a comparative methyl-differential display in ovarian tumor and normal DNA samples.

Fig. 14 depicts a Southern blot of MspI-digested genomic DNA from six different colon cancer patients which was hybridized with the BR254 probe. The number for each patient is listed above the bracket. The hybridization patterns of 6  $\mu$ g tumor (T) and 6  $\mu$ g matched normal (N) DNAs from each patient are provided for comparison. The lane labeled "probe #254" contains a clone of the isolated BR254 probe as a positive control to show that probe BR254 hybridizes strongly to itself. In the tumor DNA, the probe

hybridizes with lower and upper bands with equal intensity. In the normal DNA, the probe hybridizes mainly with the lower band. Hence, the BR254 probe detects alteration of DNA methylation in colon cancer DNA

5     DETAILED DESCRIPTION OF THE INVENTION:

The present invention provides methods and probes for detecting and isolating DNA sequences which are mutated or methylated in one tissue type but not in another. These methods are herein named Methyl- (or Mutant-) Differential Display ("MDD") methods. These MDD methods, and probes that are isolated by MDD, are useful *inter*  
10 *alia* for identifying genes that are expressed in a tissue-specific manner. Moreover, the present MDD methods and probes are useful for detecting cancerous cells, viruses, bacteria, point mutations, amplifications, deletions and genomic rearrangements.

To detect whether a DNA sequence is differentially methylated, the present invention uses methylation-sensitive restriction enzymes to cleave tester and control  
15 DNA's. These methylation-sensitive restriction enzymes will cut their DNA recognition sites when those sites are not methylated but do not cut the DNA site if it is methylated. Hence, a non-methylated tester DNA will be cut into smaller sizes than a methylated control DNA. Similarly, a hypermethylated tester DNA will not be cleaved and will give rise to larger fragments than a normally non-methylated control DNA.

20     The present invention is particularly useful for detecting and isolating DNA fragments that are normally methylated but which, for some reason, are non-methylated in a small proportion of cells. Such DNA fragments may normally be methylated for a number of reasons. For example, such DNA fragments may be normally methylated because they contain, or are associated with, genes that are rarely expressed, genes that  
25 are expressed only during early development, genes that are expressed in only certain cell-types, and the like.

While highly expressed genes are frequently unmethylated, those highly expressed genes are often not very interesting because they are "housekeeping genes" which contribute to generalized cellular functions and cellular survival, rather than to the  
30 determination of what makes one cell-type different from another. Moreover, highly expressed genes have generally already been isolated and characterized.

Unlike the housekeeping genes, non-methylated but normally-methylated, DNA fragments are present in just a small proportion of cells. Therefore, it is very difficult to detect or isolate those fragments and to distinguish them from non-methylated housekeeping genes. The present invention solves this problem by using a combination of DNA amplification and hybridization/subtraction techniques. The amplification steps enrich the genomic DNA sample with small DNA fragments that have both 5' and 3' ends cut by the methylation-sensitive enzyme. The hybridization/subtraction steps eliminate fragments encoding non-methylated housekeeping genes and fragments which are cut on only one end by the methylation-sensitive enzyme. Several rounds of amplification and hybridization/subtraction can be used if a discrete number of fragments are not isolated with the first round.

The fragments isolated by MDD are cut on both ends by the methylation-sensitive master enzyme because only fragments cut on both ends with the master enzyme are efficiently amplified. This is accomplished by using primers during DNA amplification that recognize only the master-enzyme-cut DNA ends. Such primer specificity is created by adding adaptors to the DNA ends of a sample cut with both the master enzyme and a partner restriction enzyme. The adaptors have unique sequences which hybridize with the DNA amplification primers. The partner enzyme is used to generate a mixture of DNA fragments of an appropriate size and complexity for DNA amplification but also creates a population of DNA molecules which will not accept the adaptors because they have partner-enzyme-cut ends.

Similarly, for detecting and isolating fragments which contain a mutation, DNA fragments are cut with a master enzyme and a partner (detector) enzyme, where the partner (detector) enzyme may cut normal DNA but which will not cut mutated DNA. The master enzyme normally used for MDD is a methylation-sensitive enzyme which recognizes and cuts CG-rich sites in genomic DNA. However, for detecting and isolating fragments that contain a mutation, nonmethylation sensitive master enzymes can also be used. Any restriction enzyme can be used as a master restriction enzyme for detecting and isolating mutations so long as it cuts at a different sequence than the detector enzyme. However, a master enzyme which cuts in CG-rich regions is preferred because it selects for regions associated with genes while the detector enzyme identifies and permits isolation of the mutation by cutting only the nonmutant DNA fragments

To confirm whether fragments isolated by the present MDD methods can detect differentially methylated or mutated DNA sites, they are tested as probes on genomic DNAs which are cut with methylation-sensitive or mutant-detector enzymes, respectively. If the probe detects a different size fragment in DNA which is known, or suspected to be, hypomethylated, hypermethylated or mutated than in normal DNA, then the probe can detect differentially methylated or mutated DNA sites.

DNA fragments isolated by the present methods can therefore constitute probes, for example, for detecting hypomethylation, hypermethylation or mutation in DNA fragments which are normally methylated.

In particular, the present invention provides a method of isolating probes to detect hypomethylation in a CNGtriplet of DNA which includes:

- a) cleaving a tester sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;
- b) amplifying the cleaved tester sample using primers that recognize DNA ends cut by the master restriction enzyme, to generate a first tester amplicon;
- c) melting and hybridizing the first test amplicon with a driver DNA to generate a tester-tester product and a tester-driver product;
- d) adding nucleotides to the ends of the tester-tester product and the tester-driver product to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;
- e) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by using primers which recognize only master-enzyme-cut DNA ends, to produce a second tester amplicon;
- f) isolating a discrete DNA fragment from the second tester amplicon as a probe to detect hypomethylation a CNGtriplet of DNA;

wherein the master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence;

wherein the partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and



wherein the driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by the master restriction enzyme.

5 In another embodiment, the master-enzyme-cut DNA ends are distinguished from the partner-enzyme-cut DNA ends by ligating a first set of adaptors onto master-enzyme-cut DNA ends of the cleaved tester sample to generate a first-tester amplification template. This first-tester amplification template is then amplified to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to the first set of adaptors. Hence, only DNA fragments which were cut by the master restriction  
10 enzyme will be amplified. The first adaptors are then cleaved off the first-tester amplicon and a second set of adaptors is ligated onto master-enzyme-cut DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends. As before, this tester DNA is melted and hybridized with a large molar excess of a driver DNA to generate a mixture of tester-tester product and tester-driver product. The DNA  
15 ends of these products are made blunt by adding nucleotides and the blunt-ended tester-driver product is amplified the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends. This generates a second-tester amplicon from which a discrete DNA fragments can be isolated as probes to detect the hypomethylation in a CNGtriplet of  
20 DNA.

As used herein, hypomethylation means that at least one cytosine in a CG or CNGdi- or tri-nucleotide site in genomic DNA of a given cell-type does not contain CH<sub>3</sub> at the fifth position of the cytosine base. Cell types which may have hypomethylated CGs or CCGs include any cell type which may be expressing a non-housekeeping function.  
25 This includes both normal cells that express tissue-specific or cell-type specific genetic functions, as well as tumorous, cancerous, and similar cell types. Cancerous cell types and conditions which can be analyzed, diagnosed or used to obtaining probes by the present methods include Wilm's cancer, breast cancer, ovarian cancer, colon cancer, kidney cell cancer, liver cell cancer, lung cancer, leukemia, rhabdomyosarcoma, sarcoma, and  
30 hepatoblastoma.

Genomic DNA samples can be obtained from any mammalian body fluid, secretion, cell-type, or tissue, as well as any cultured cell or tissue.

The present MDD technique involves preparation of a driver DNA and a tester DNA. A driver DNA is genomic DNA from a "normal" cell type. A tester DNA is genomic DNA from a "test" cell type which may have mutated DNA sites, hypermethylated DNA sites or non-methylated DNA sites in place of the normal DNA at that site. Such normal cell types are cells of the same tissue-type as the "test" cell type, but which are normally methylated and are not mutated. Test cell types include cells that express tissue or cell-specific functions, cancer cells, tumor cells and the like. For example, if breast cancer cells are selected as the test cell type, the normal cell type would be breast cells that are not diseased or cancerous. Preferably, test and normal cell types are isolated from the same person.

After isolation, the tester DNA and the driver DNA are separately cut with both the master and partner restriction enzymes. As provided by the present invention, the master restriction enzyme cleaves a CNG DNA sequence. The master restriction enzyme produces sticky ends, rather than blunt ends, when it cuts. In general, the master enzyme used in the present methods cleaves only non-methylated CNG DNA sequences and does not cleave methylated CNG DNA sequences. However, for selected applications a non-methylation sensitive master enzyme can be used, for example, for detecting and isolating mutations. For detecting and isolating hypomethylated DNA sites, a methylation-sensitive master enzyme should be used, for example, the MspI restriction enzyme which recognizes and cleaves DNA at nonmethylated CCGG but will not cleave the CCGG sequence when the outer cytosine is methylated. Master restriction enzymes contemplated by the present invention include MspI, BsiSI, HinfI and the like. In a preferred embodiment, MspI is used as the master restriction enzyme for detecting and isolating a DNA fragment containing a hypomethylated site.

According to the present invention, a partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme. As used herein, the complexity of mixture of DNA fragments is the number of different DNA fragments within a selected size range. When mammalian DNA is used, the present methods can effectively screen a DNA sample containing about 1% to about 25% of the mammalian genome. In a preferred embodiment, a DNA sample has about 5% to about 15% of the total DNA sequences in the mammalian genome. In an especially preferred

embodiment, the DNA sample has about 10% of the total DNA sequences in the mammalian genome.

According to the present invention, the size range of DNA fragments in a sample of genomic DNA is that size range which can be amplified by in vitro DNA amplification. When polymerase chain reaction ("PCR") is used, that size range is about 2000 base pairs (bp) to about 75 bp. Preferably the selected size range is about 1500 bp to about 100 bp. More preferably, the selected size is about 1000 bp to about 150 bp.

As used herein, the preferred size is generated by the action of both the master and partner restriction enzymes. Hence, if the partner restriction enzyme generates a population of genomic DNA fragments which are about 1000 bp or smaller, but, in combination with the master restriction enzyme a population of fragments which are less than 50 bp is generated, a different set of partner and master enzymes is preferably used.

Also according to the present invention, the partner restriction enzyme cleaves DNA to produce ends that are neither homologous nor complementary to a sticky end produced by the master restriction enzyme.

Any convenient restriction enzyme is used as a partner restriction enzyme to identify CNGsites that are hypomethylated, so long as the appropriate complexity and size of DNA fragments are generated. For convenience, a partner restriction enzyme that operates under the same conditions as the master restriction enzyme may be used. For example, MseI is a good partner restriction enzyme which can be used under the same conditions as the preferred MspI master restriction enzyme. When the preferred MspI master restriction enzyme is used, the partner restriction enzyme preferably recognizes and cuts a 4-nucleotide sequence. Moreover, in a preferred embodiment, the partner restriction enzyme cuts an AT-rich site or generates a blunt end to avoid any ligation of adaptors, which are designed to attach to master enzyme cut ends, onto the ends generated by the partner enzyme. Preferred partner restriction enzymes include MseI, Sau3A, RsaI, TspEI, MaeI, NlaIII, DpnI and the like.

Cleavage methods and procedures for selecting restriction enzymes for cutting DNA at specific sites are known to the skilled artisan. For example, many suppliers of restriction enzymes provide information on conditions and types of DNA sequences cut by specific restriction enzymes, including New England BioLabs, Pro-Mega Biochems, Boehringer-Mannheim and the like. Sambrook et al. provide a general description of

methods for using restriction enzymes and other enzymes. See Sambrook et al., 1989 Molecular Cloning: A Laboratory Manual, Vols. 1-3 (Cold Spring Harbor Press, Cold Spring Harbor, NY).

After cleaving the DNA with the master and partner restriction enzymes, a set of  
5 adaptors is hybridized and ligated onto the sticky DNA ends produced by the master  
restriction enzyme. The adaptors are selected so that they will ligate to the CG-rich ends  
of the DNA cut by the master restriction enzyme but not to the ends of the fragments that  
were cut with the partner restriction enzyme. Because the only the DNA ends of  
fragments in the tester and driver genomic DNA samples will be single-stranded, adaptors  
10 need not be longer than about 12 nucleotides to specifically recognize and hybridize to the  
master enzyme-cut ends. However, the adaptors are chosen not only to ligate to DNA  
ends cut by the master restriction enzyme, but also to be a good size and DNA sequence  
to act as recognition sites for primers used during DNA amplification. Adaptors are  
chosen so that, when the mixture of master and partner enzyme cut DNA fragments is  
15 amplified using primers that hybridize to those adaptors, only those fragments that have  
the adaptor and thus were cut with the master restriction enzyme will be amplified.  
Hence, after hybridization and ligation the adaptor ends should have a unique sequence  
and be of sufficient length to be a unique recognition site for primers selected for in vitro  
DNA amplification.

20 Preferred sets of adaptors of the present invention include those of SEQ ID NO.  
1-6. As used herein:

the MSA24 adaptor has SEQ ID NO:1 -- CTCGTCGTCAGGTCAGTGCTTCAC

the MSA12 adaptor has SEQ ID NO:2 -- CGGTGAAGCACT

the MSB24 adaptor has SEQ ID NO:3 -- TAGAGCCACGTAGCTGCTGTAGTC

25 the MSB12 adaptor has SEQ ID NO:4 -- CGGACTACAGCA

the MSC24 adaptor has SEQ ID NO:5 -- ACCGTGGACTGGATAGGTTTCAGAC

and the MSC12 adaptor has SEQ ID NO:6 -- CGGTCTGAACCT

In one embodiment, the SEQ ID NOS: 1 and 2 (MSA24 and MSA12) adaptors  
are single-stranded and are used as the first set of adaptors. These single-stranded  
30 adaptors are annealed and hybridized onto DNA ends cut with the master restriction  
enzyme and then ligated to those DNA ends by incubation with ligase. Both the tester and  
driver DNAs receive the first set of adaptors. Thus, the master-cut ends of both the tester

and driver DNAs will have adaptors while DNA ends cut with the partner enzyme, from either the tester or driver, will have no such adaptors.

The two samples are then separately amplified by known DNA amplification procedures using primers that hybridize to the adaptors. For example, when the SEQ ID NO: 1 and 2 adaptors are used, an oligonucleotide having SEQ ID NO: 1 can be used as a primer for DNA amplification of only the master enzyme cut DNA fragments. This generates a tester amplicon and a driver amplicon. In both the tester and driver amplicons, only fragments which are cut with the master restriction enzyme will be amplified.

DNA amplification procedures for use in the present MDD methods include any in vitro amplification procedure which provides sequence-specific synthesis of a nucleic acid fragment relative to the complex bulk of nucleic acid present in a sample. The specificity of the process is determined by enzymatic recognition of a specific sequence or by the oligonucleotide primers which are capable of hybridizing only with selected adaptors that are added onto the ends of master restriction cut DNA fragments.

In vitro DNA amplification techniques are known in the art. A review of such techniques can be found in Kwok et al., *8 Am. Biotechnol. Lab.* 14 (1990). In vitro nucleic acid amplification techniques include polymerase chain reaction (PCR), transcription-based amplification system (TAS), self-sustained sequence replication system (SSR), ligation amplification reaction (LAR), ligase-based amplification system (LAS), Q $\beta$  RNA replication system and run-off transcription.

PCR is a preferred method for DNA amplification. PCR synthesis of DNA fragments occurs by repeated cycles of heat denaturation of DNA fragments, primer annealing onto the adaptor ends of the master-cut DNA fragment, and primer extension. These cycles can be performed manually or, preferably, automatically. Thermal cyclers such as the Perkin-Elmer Cetus cycler are specifically designed for automating the PCR process, and are preferred. The number of cycles per round of synthesis can be varied from 2 to more than 50, and is readily determined by considering the source and amount of the nucleic acid template, the desired yield and the procedure for detection of the synthesized DNA fragment.

PCR techniques and many variations of PCR are known. Basic PCR techniques are described by Saiki *et al.* (1988 Science 239:487-491) and by U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, which are incorporated herein by reference.

5 The conditions generally required for PCR include temperature, salt, cation, pH and related conditions needed for efficient copying of the master-cut fragment. PCR conditions include repeated cycles of heat denaturation (i.e. heating to at least about 95°C) and incubation at a temperature permitting primer: adaptor hybridization and copying of the master-cut DNA fragment by the amplification enzyme. Heat stable amplification enzymes like the pwo, Thermus aquaticus or Thermococcus litoralis DNA  
10 polymerases are commercially available which eliminate the need to add enzyme after each denaturation cycle. The salt, cation, pH and related factors needed for enzymatic amplification activity are available from commercial manufacturers of amplification enzymes.

As provided herein an amplification enzyme is any enzyme which can be used for  
15 in vitro nucleic acid amplification, e.g. by the above-described procedures. Such amplification enzymes include pwo, Escherichia coli DNA polymerase I, Klenow fragment of E. coli DNA polymerase I, T4 DNA polymerase, T7 DNA polymerase, Thermus aquaticus (Taq) DNA polymerase, Thermococcus litoralis DNA polymerase, SP6 RNA polymerase, T7 RNA polymerase, T3 RNA polymerase, T4 polynucleotide kinase, Avian  
20 Myeloblastosis Virus reverse transcriptase, Moloney Murine Leukemia Virus reverse transcriptase, T4 DNA ligase, E. coli DNA ligase or Q $\beta$  replicase. Preferred amplification enzymes are the pwo and Taq polymerases. The pwo enzyme is especially preferred because of its fidelity in replicating DNA.

The length of the primers for use in MDD depends on several factors including the  
25 nucleotide sequence and the temperature at which these nucleic acids are hybridized or used during in vitro nucleic acid amplification. The considerations necessary to determine a preferred length for an oligonucleotide primer are well known to the skilled artisan. For example, as is known to the skilled artisan, the length of a short nucleic acid or oligonucleotide can relate to its hybridization specificity or selectivity. Because the tester  
30 and driver samples contain a complex mixtures of nucleic acids, primers which are shorter than about 12 nucleotides may hybridize to more than one site in the test genomic DNA, and accordingly would not have sufficient hybridization selectivity for amplifying only the

master-cut fragments. However, a 12- to 15-nucleotide sequence is generally represented only once in a mammalian genome (Sambrook *et al.* 1989 Molecular Cloning: A Laboratory Manual, Vol. 2, Cold Spring Harbor Press, Cold Spring Harbor, NY; pp 11.7-11.8). Accordingly, to eliminate amplification of fragments which do not have  
5 adaptors, primers are chosen which are generally at least about 14 nucleotides long.

Preferably, the present primers are at least 16 nucleotides in length. More preferred primers are at least 17 nucleotides in length (Sambrook *et al.*, pp. 11.7- 11.8).

After amplification, the first set of adaptors are removed from the ends of the tester and driver amplicons by cleavage with the master restriction enzyme. Cleavage with  
10 the master restriction enzyme preserves the original ends of the fragments and removes all DNA which was not present in the tester DNA as originally isolated.

A second set of adaptors are ligated to the tester amplicon, but not to the driver amplicon. This generates a second-adaptor-tester sample. The second set of adaptors is selected so that they do not have the same sequence as the first set of adaptors and so that  
15 they hybridize, and hence ligate, only to DNA ends cut by the master restriction enzyme. The second set of adaptors should also form a good recognition site for primers which are used during DNA amplification

During DNA amplification, non-methylated DNA fragments which are normally present in all cell types, and which may not be of interest, will also be amplified because  
20 they will have adaptors on their ends. At least one round of subtraction/hybridization followed by DNA amplification is used by the present MDD methods to remove these nonmethylated DNA fragments which likely encode "housekeeping" functions.

In general, the present MDD techniques remove "housekeeping" sequences by hybridizing the amplified tester DNA with a large molar excess of amplified driver DNA and amplifying only the tester-tester hybrids which have perfectly hybridized ends because  
25 those perfectly hybridized ends are the only sites recognized by the primers used for DNA amplification.

In particular, the amplified tester DNA (called the first test amplicon or the second-adaptor-tester) is mixed with a large molar excess of amplified driver DNA. The  
30 mixture is denatured, then renatured. As used herein, a large molar excess of driver to tester DNA is about a 10-fold to about a 10,000-fold molar excess of driver DNA to tester DNA. In a preferred embodiment the tester DNA is melted and hybridized with

about a 10-fold to about a 1000-fold molar excess of a driver DNA. In a more preferred embodiment, the tester DNA is melted and hybridized with about a 100-fold molar excess of a driver DNA. Stringent hybridization conditions are preferably used during renaturation to maximize the formation of fragments with few mismatches.

5           The skilled artisan can vary the hybridization conditions used during MDD to achieve optimal hybridization of unique tester DNA fragments to their tester homologue. Conditions for achieving optimal hybridization are known to the skilled artisan and generally include temperature and salt concentrations permitting selective hybridization between two highly homologous DNA fragments, e.g. stringent hybridization conditions. Stringent conditions permit little or no detectable hybridization between mismatched driver or tester fragments, that is between fragments that have dissimilar sequences, particularly at the ends. Hybridization techniques are described, for example, in

10           Sambrook et al., 1989 Molecular Cloning: A Laboratory Manual, Vols. 1-3 (Cold Spring Harbor Press, Cold Spring Harbor, NY).

15           The ends of the renatured fragments are then filled in using a DNA polymerase, e.g. Taq or pwo polymerase. Such a "filling in" reaction makes the DNA fragments blunt-ended but does not eliminate any internally mismatched or unhybridized regions. To eliminate such mismatched hybrids, the renatured fragments may be treated with a nuclease that will remove single-stranded regions and mismatched nucleotides in the renatured fragments. For example, S1 nuclease, mung bean nuclease and the like can be

20           used to eliminate mismatched hybrids. The nuclease will also cleave off mismatched nucleotides at the ends of either the tester or the driver fragment. Hence, mismatched hybrids will have ends that do not contain second adaptor ends. Only the single-stranded tester fragments with second adaptor ends that faithfully hybridize to an exact tester

25           homologue will end up with a second adaptor end.

          The subtraction/hybridization mixture is then amplified by in vitro DNA amplification procedures using primers that hybridize to the second adaptor ends. Hence, only the tester DNA fragments with second adaptor ends are amplified. Any tester DNA which hybridized with driver DNA will not be amplified. A large excess of driver DNA

30           is used to promote formation of hybrids that are commonly found in both the tester and driver samples. The result is a second tester amplicon with fragments which were uniquely unmethylated in the original tester DNA sample.



DNA fragments cut on only one end with the master restriction enzyme ("single master-cut fragments") are eliminated by the present invention in several ways. First, only one end of those fragments will ligate to an adaptor. This means that, during any amplification step, the single master-cut fragments will not be amplified as quickly as the "double master-cut fragments" which have adaptors at both ends. As is known to the skilled artisan, fragments which hybridize with amplification primers at both ends are amplified geometrically, while DNA fragments which hybridize with amplification primers at only one end are amplified arithmetically. Second, during the subtraction/hybridization step those single master-cut fragments may also hybridize to their driver homologues because the driver is present in vast molar excess. This imperfect hybrid can be eliminated by treatment with nuclease, so it will not be amplified during the subsequent DNA amplification step.

Several rounds of amplification and subtraction/hybridization can be used to isolate discrete DNA fragments which can be used as probes to detect hypomethylation, hypermethylation and mutation in specific regions of the genome. Discrete DNA fragments isolated by the present MDD methods which can be used to detect hypomethylation, hypermethylation and mutations in genomic DNA samples from different cell types include DNA fragments having SEQ ID NO: 7-10.

While discrete DNA fragments which detect hypomethylation also frequently detect hypermethylation, the MDD method is readily adapted to permit isolation of fragments which will specifically detect hypermethylation in identified sites of the genome. This is accomplished merely by using normal cell DNA as the source of DNA for the hypermethylation probe. In this case, the normally unmethylated DNA will be cut by the methylation-sensitive master enzyme while the hypermethylated test DNA will not be cut. To eliminate DNA fragments which are common to the normal and test DNA samples, the normal cell DNA is hybridized with an excess of test cell DNA (now called "test driver DNA"). Hence, isolation of hypermethylation probes generally involves the same manipulations, as used for isolation of hypomethylation probes, but the roles of the normal cell DNA and the test cell DNA are reversed.

Also for detecting and isolating DNA fragments that have a hypermethylated CNG, a methylation-sensitive enzyme can be used as a partner enzyme and a non-

methylation sensitive enzyme as a master enzyme. When this modification is used, the adaptors are still designed to attach to the ends of DNA fragments cut by the non-methylation sensitive master restriction enzyme.

5 In particular, the present invention provides a method of isolating a probe to detect hypermethylation in a CNGtriplet of DNA which includes:

- a) cleaving a normal sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved normal sample;
- b) amplifying the cleaved normal sample using primers that recognize DNA ends cut by the master restriction enzyme, to generate a first normal amplicon;
- 10 c) melting and hybridizing the first normal amplicon with a test driver DNA to generate a normal-normal product and a normal-driver product;
- d) adding nucleotides to the ends of the normal-normal product and the normal-driver product to make a blunt-ended normal-normal product and a blunt-ended normal-driver product;
- 15 e) amplifying the blunt-ended normal-normal product and the blunt-ended normal-driver product by using primers which recognize only master-enzyme-cut DNA ends, to produce a second normal amplicon;
- f) isolating a discrete DNA fragment from the second normal amplicon as a probe to detect the hypermethylation in a CNGtriplet of DNA;

20 wherein the test driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by the master restriction enzyme.

The normal sample of genomic DNA is obtained by isolating genomic DNA from a "normal" cell type, and the test driver DNA is obtained by isolating genomic DNA from  
25 a "test" cell type which may have hypermethylated DNA sites in place of the normal non-methylated DNA at that site. Here, the test cell type is a cancerous or tumorous cell type.

In a further embodiment, the present MDD techniques can be used to identify and isolate mutations in genomic DNA and it has particular utility for identifying mutations which are near the promoters or coding regions of genes. Nonmethylation-sensitive  
30 master restriction enzymes can be used for isolating mutations. However, the preferred master enzymes recognize CNG DNA sequences including CpG dinucleotides which are frequently methylated in human genome. Condensed CpG islands appear in gene

promoter regions, or even in the first exons. Because CpG islands appear in gene promoter regions, and in the first few exons of many genes, MDD-identified lesions will often be close to genes which cause disease.

5 According to the present invention, MDD can be used to isolate probes for detecting any type of genomic lesion. For example, MDD can be used to identify fragments which have point mutations, deletions, insertions, amplifications, rearrangements, and other mutations.

10 MDD is readily adapted for isolating mutations in DNA sequence by using a partner enzyme, now called a detector enzyme, that cuts normal DNA but which will not cut the mutated DNA. For example, a point mutation in tester DNA can be identified when a cleavage site recognized by the detector enzyme is eliminated. When cut with master and detector restriction enzymes, a fragment with both ends cleaved by the master enzyme will be present in the tester DNA. However, the detector enzyme cleaves this fragment in the driver DNA, because normal DNA has no point mutation. As a result, the  
15 fragment containing the point mutation will be isolated after performing the present MDD amplification and subtraction/hybridization steps. The point mutation can be easily identified by comparing the DNA sequence of the isolated mutant fragment with the DNA sequence of the homologue isolated from a matched normal genomic DNA. Moreover, the mutant tester fragment will contain the entire mutation -- rather than just part of the  
20 mutation as would occur if a detector enzyme were used that cut the mutation rather than the normal DNA site.

Similarly, MDD is readily adapted for isolating deletion mutations in DNA ismply by using tester DNA as the driver DNA and normal DNA as the source of the intact fragment whose homologue in the tester DNA has a deletion. The deletion mutant  
25 fragment can be isolated by virtue of its partial homology with the normal DNA fragment.

In particular, the present invention provides a method of identifying a probe to detect a mutation in a tester sample of genomic DNA which includes:

- a) cleaving a tester sample of genomic DNA with both a master restriction enzyme and a detector restriction enzyme to generate a cleaved tester sample;
- 30 b) ligating a first set of adaptors onto master enzyme cut DNA ends of the cleaved tester sample to generate a first-tester amplification template;

c) amplifying the first-tester amplification template to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to the first set of adaptors;

5 d) cleaving off the first adaptors from the first-tester amplicon and ligating a second set of adaptors onto DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends;

e) melting and hybridizing the second-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a mixture of tester-tester product and tester-driver product;

10 f) adding nucleotides onto DNA ends present in the mixture to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

g) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-tester amplicon;

15 h) isolating a discrete DNA fragment from the second tester amplicon as a probe to detect a mutation in the tester sample of genomic DNA;

20 wherein the detector restriction enzyme cleaves a normal DNA site but a mutant DNA site to produce DNA fragments with a complexity of about 5% to about 25% of the genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein the driver DNA is cut with both the master restriction enzyme and the detector restriction enzyme and amplified using primers that recognize DNA ends cut by the master restriction enzyme.

25 Probes isolated by the present MDD techniques have at least about 14 nucleotides to about 2000 nucleotides.

30 For identifying mutations and isolating DNA fragments containing those mutations, selection of a detector restriction enzyme involves identifying an enzyme which will not cut mutated genomic DNA but which will cut normal DNA. The skilled artisan can readily perform this analysis, for example, by testing a plurality of restriction enzymes in the present MDD methods.

According to the present invention, a mutation, hypomethylation and hypermethylation at a specific site in genomic DNA can be detected by observing whether

the size or intensity of a DNA fragment cut with a master or detector restriction enzyme is the same in control and test samples. This can be done by cutting genomic DNA isolated from control and test tissue samples with the master or detector restriction enzyme, hybridizing a probe to the control and test DNAs and observing whether the two hybridization complexes are the same or different sizes or intensities.

In particular, the present invention provides a method of detecting whether a CNGtriplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cells which includes:

- a) isolating genomic DNA from a control sample of cells and a test sample of cells to generate a control-cell DNA and a test-cell DNA;
- b) cleaving the control-cell DNA and the test-cell DNA with a master restriction enzyme to generate cleaved control-cell DNA and cleaved test-cell DNA;
- c) preparing a probe from a DNA isolated by the methods of the present invention, for example, a DNA selected from the group consisting of SEQ ID NO: 7-10;
- d) hybridizing the probe to the cleaved control-cell DNA and the cleaved test-cell DNA to form a control-hybridization complex and a test-hybridization complex; and
- e) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex.

Similarly, the present invention provides a method of detecting whether a DNA site is mutated in a genomic DNA present in a test sample of cells which includes:

- a) isolating genomic DNA from a control sample of cells and a test sample of cells to generate a control-cell DNA and a test-cell DNA;
- b) cleaving the control-cell DNA and the test-cell DNA with a detector restriction enzyme to generate cleaved control-cell DNA and cleaved test-cell DNA;
- c) preparing a probe from a discrete DNA isolated by the methods of the present invention;
- d) hybridizing the probe to the cleaved control-cell DNA and the cleaved test-cell DNA to form a control-hybridization complex and a test-hybridization complex; and
- e) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex;

wherein the detector restriction enzyme does not cleave a mutated DNA site but does cleave a corresponding nonmutated DNA site.

Hybridization techniques for detecting specific genomic DNA fragments are known in the art and include solid-phase-based hybridization and solution hybridization which use any of the known reporter molecules. Detailed methodology for gel electrophoretic and nucleic acid hybridization techniques can be found in Sambrook et al.

- 5 Preferred nucleic acid probes for detecting hypomethylation and hypermethylation have SEQ ID NOS:7-10.

According to the present invention, a probe isolated by the subject methods, can be labelled by any procedure known in the art, for example by incorporation of nucleotides linked to a "reporter molecule".

- 10 A "reporter molecule", as used herein, is a molecule which provides an analytically identifiable signal allowing detection of a hybridized probe. Detection may be either qualitative or quantitative. Commonly used reporter molecules include fluorophores, enzymes, biotin, chemiluminescent molecules, bioluminescent molecules, digoxigenin, avidin, streptavidin or radioisotopes. Commonly used enzymes include horseradish  
15 peroxidase, alkaline phosphatase, glucose oxidase and  $\beta$ -galactosidase, among others. Enzymes can be conjugated to avidin or streptavidin for use with a biotinylated probe. Similarly, probes can be conjugated to avidin or streptavidin for use with a biotinylated enzyme. The substrates to be used with these enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change.  
20 For example, p-nitrophenyl phosphate is suitable for use with alkaline phosphatase reporter molecules; for horseradish peroxidase, 1,2-phenylenediamine, 5-aminosalicylic acid or toluidine are commonly used.

- Incorporation of a reporter molecule into a DNA probe can be by any method known to the skilled artisan, for example by nick translation, primer extension, random  
25 oligo priming, by 3' or 5' end labeling or by other means (see, for example, Sambrook et al.).

- In another embodiment, the present invention provides one or more compartmentalized kits for detection of hypomethylation. A first kit has a receptacle containing at least one of the present isolated probes. Such a probe may be a nucleic acid  
30 fragment which is methylated at a CNG trinucleotide in the genomic DNA of normal cells but which not methylated in the genomic DNA of cancerous cells. Such a probe may be specific for a DNA site that is normally methylated but which is nonmethylated in certain

cell types. Similarly, such a probe may be specific for a DNA site that is abnormally methylated or hypermethylated in certain cell types. Finally, such a probe may identify a specific DNA mutation. The probe provided in the present kits may have a covalently attached reporter molecule.

5           A second kit has at least one receptacle containing a reagent for MDD. Such a reagent can be, for example, a preferred master restriction enzyme, a preferred partner restriction enzyme, a preferred adaptor which can be ligated to DNA cut with the preferred master restriction enzyme, a primer that hybridizes to a preferred adaptor, and the like.

10           Using the MDD methods of the present invention, several methyl-polymorphic markers or sites in the human genome have been identified and DNA fragments encoding these marker sites have been isolated from different cell types. In particular, the present MDD methods have been used to identify hypomethylated sites which act as markers indicating whether a patient has cancer. These sites were identified using B-cells from  
15 patients with Chronic Lymphocytic Leukemia (CLL) and using tumor cells from breast cancer patients.

Moreover, for the first time the present methods and probes experimentally prove that methylated external cytosine at CpCpG triplet sequences do exist in the human genome, and that alteration of their methylation patterns is related to human cancer  
20 development. According to the present invention, methylation of the first base of CpCpG triplets is not a random event. Instead, it is a site- and tissue-specific event.

The alteration of the methylation pattern in CpCpG triplets may be a key, and common event, in the development of neoplasia. Imbalanced methylation at CpCpG triplet may play at least two roles in tumorigenesis:

25           1) DNA hypomethylation may cause some proto-oncogene expression or DNA hypermethylation may silence a tumor suppressor which contributes to neoplastic growth; and

30           2) DNA hypomethylation may change chromatin structure, and induce abnormalities in chromosome pairing and disjunction. Such structural abnormalities may result in genomic lesions, such as chromosome deletions, amplifications, inversions, mutations, and translocations, all of which are found in human genetic diseases and cancer.

Hence, the present MDD methods have broad utility for identifying differentially methylated sites at CpCpG triplet sequences in the human genome; for mapping hypo- and hyper-methylation sites which are related to disease development; for understanding the role(s) of DNA methylation in normal cell genomic DNA imprinting, differentiation, and development; for understanding tumorigenesis; for diagnosing and monitoring the prognosis of disease; and for searching for proto-oncogene(s) and recessive oncogene(s). MDD, as a comprehensive molecular genetic technique, enables the accumulation of knowledge necessary to help better understand the etiology of human genetic diseases and cancer.

The following examples further illustrate the invention.



### Example 1

#### Materials and Methods

To test the MDD technique, a working model system was designed to isolate methyl-polymorphic DNA markers (see Fig. 1 for a schematic diagram of the MDD methods). There is evidence that in mammalian DNA, methylation patterns are both tissue-, and cell-type specific. Hence, DNA was isolated from two different types of cells from patients who have chronic lymphocytic leukemia (CLL). In CLL patients, the number of CD5<sup>+</sup> B type lymphocytes is abnormally high. These CD5<sup>+</sup> cells can be isolated by highly specific antibodies in a fluorescence activated cell sorter (FACs). For performing MDD, the DNA isolated from these CD5<sup>+</sup> cells served as the tester, while DNA isolated from neutrophils of the same patient served as the driver. A methyl-polymorphic DNA probe, named CLL58, which exhibited methyl-differential displays was isolated from an individual CLL patient.

After the successful isolation of these methyl-polymorphic probes, MDD was again successfully used in the isolation of methylpolymorphic probes from breast cancer biopsies. Three methyl-polymorphic probes, BR50, BR104, and BR254, which gave methyl-differential displays in breast cancer and their matched normal cells, were isolated from three individual patients. BR50 and BR254 also gave methyl-differential displays in ovarian cancer and their matched normal cells, while the BR254 probe exhibited a methyl-differential display in colon cancer cells.

DNA Isolation from malignant B-cells and neutrophil cells. Mononuclear and neutrophil cells were isolated from the peripheral whole blood of CLL patients by Ficoll/Hypaque centrifugation. Leukemic B-cells (CD5<sup>+</sup>), were sorted from mononuclear cells in the presence of fluorescent anti-CD5 antibodies by FACs. DNAs were isolated by the phenol extraction method. Sambrook, J., Fritsch, E. F. and Manitis, T., (1989) Molecular Cloning, 9.16-9.19.

DNA from human cancer biopsies. Dissected human breast, ovarian, and colon cancer tissues (tumor and normal) were immediately frozen in liquid nitrogen, and stored at -70°C. DNAs were isolated from both tumor and matched normal tissues by the phenol

extraction method. Sambrook, J., Fritsch, E. F. and Manitis, T., (1989) Molecular Cloning, 9.16-9.19.

5       **MDD.** To prepare tester and driver amplicons, 1-2  $\mu\text{g}$  of tester and driver genomic DNA were digested simultaneously with Msp I and MseI restriction endonucleases (New England Biolabs). The digested DNAs were extracted with phenol/chloroform once, and precipitated with 2 volumes of ethanol after adding 1/10 volume of 3 M NaAc.

10       The DNAs were then resuspended in 10  $\mu\text{l}$  of TE buffer (10 mM Tris.HCl, 0.1 mM EDTA, pH8.0). To ligate the first pair of oligonucleotide adapters (MSA24 with SEQ ID NO:1 and MSA12 with SEQ ID NO:2) to the digested tester and driver amplicon DNAs, 10  $\mu\text{l}$  (1  $\mu\text{g}$ ) of digested DNA was mixed with: 2.5  $\mu\text{l}$  of each of the unphosphorylated single strand adapters having SEQ ID NO:1 (0.6  $\mu\text{g}/\mu\text{l}$ ) and SEQ ID NO:2 (0.3  $\mu\text{g}/\mu\text{l}$ ); 2  $\mu\text{l}$  of 10X ligation buffer (Boehringer Mannheim); and 3  $\mu\text{l}$  of ddH<sub>2</sub>O. The oligonucleotide adapters were annealed by cooling the mix gradually from 15       55°C to 10°C over a 1.5 hour period in a 4°C cold room, and then ligated to MspI/MseI-cut DNA fragments by overnight incubation with 1 U of T4 DNA ligase (Boehringer Mannheim) at 16°C. After ligation, the ligate was diluted with TE buffer (10 mM Tris.HCl, 0.1 mM EDTA, 20  $\mu\text{g}/\text{ml}$  of tRNA) at a concentration of 2  $\text{ng}/\mu\text{l}$ .

20       To generate the tester and driver amplicons, tester and driver DNA ligated to the first set adapters were separately amplified by polymerase chain reaction ("PCR"). Two PCR tubes were set up to make the tester amplicon, while five PCR tubes were prepared for the driver amplicon. Each tube has a 400  $\mu\text{l}$  reaction mixture containing 67 mM Tris-HCl, pH 8.8 at 25°C, 4 mM MgCl<sub>2</sub>, 16 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM  $\beta$ -mercaptoethanol, 100  $\mu\text{g}/\text{ml}$  of non-acetylated bovine serum albumin, 24  $\mu\text{l}$  of dNTP (5 mM for each dATP, 25       dCTP, dGTP, and dTTP), 8  $\mu\text{l}$  of SEQ ID NO:1 (0.6  $\mu\text{g}/\mu\text{l}$ ), and 80 ng of DNA ligated to the adapters as templates. While the tubes were equilibrated at 72°C in a thermal cycler (Perkin-Elmer Cetus), 12U of Taq DNA polymerase (Perkin-Elmer Cetus), and 3U of pwo DNA polymerase (Boehringer Mannheim) were added. The reaction was overlaid with mineral oil and incubated for 5 minutes to fill-in the 3'-recessed ends of the ligates, 30       and the DNAs were then amplified by 24 cycles of PCR. Each PCR cycle included 1 minute at 95°C, 1 minute at 67°C and 3 minutes at 72°C, and the last cycle was followed with a 10 minute extension time at 72°C. The DNA amplicons were extracted once with

phenol, once with phenol/chloroform, and precipitated by 2 volumes of 100% ethanol after adding 1/4 volume of 10 M  $\text{NH}_4\text{Ac}$ . Amplicon DNAs were then resuspended in TE buffer.

5 To remove the first set of adapters from the driver amplicon, 50  $\mu\text{g}$  of driver amplicon DNA was digested to completion with *Msp* I enzyme (10U/ $\mu\text{g}$ ). The driver amplicon DNA was extracted once with phenol/chloroform, and after adding 1/4 volume of 10 M  $\text{NH}_4\text{Ac}$ , the DNA was precipitated with 2 volumes of 100% ethanol. The amplicon DNA was resuspended in TE buffer at a concentration of 0.5  $\mu\text{g}/\mu\text{l}$ . To change the tester amplicon DNA adapters, 5  $\mu\text{g}$  of tester amplicon was digested with *Msp* I  
10 (10U/ $\mu\text{g}$ ). The digested tester amplicon DNA was purified by PCR purification spin column (Boehringer Mannheim).

For the hybridization/subtraction and amplification steps, *Msp* I digested tester amplicon DNA was ligated to a second set of adapters (MSB24 having SEQ ID NO:3 and MSB12 having SEQ ID NO:4): 1  $\mu\text{l}$  (0.2  $\mu\text{g}$ ) of digested tester DNA was mixed with 1  
15  $\mu\text{l}$  of the SEQ ID NO:3 adaptor (0.6  $\mu\text{g}/\mu\text{l}$ ), 1  $\mu\text{l}$  of the SEQ ID NO:4 adaptor (0.3  $\mu\text{g}/\mu\text{l}$ ), 1  $\mu\text{l}$  of 10X ligation buffer, and 6  $\mu\text{l}$  of  $\text{ddH}_2\text{O}$ . The same annealing and ligation steps were performed as for ligation of the first set of adapters. After ligation, the ligate was diluted with 40  $\mu\text{l}$  of TE buffer and mixed with 20  $\mu\text{g}$  (40  $\mu\text{l}$ ) of driver amplicon DNA. The mix was extracted with phenol/chloroform once, and precipitated by adding  
20 1/4 volume of 10 M  $\text{NH}_4\text{Ac}$ , and 2.5 volumes of 100% ethanol in a dry ice/ethanol bath for 10 minutes. The reaction solution was equilibrated to room temperature and the DNA was collected by centrifugation. The pellet was washed twice with 70% ice cold ethanol, dried, resuspended in 2  $\mu\text{l}$  of 3XEE buffer. D. Straus et al., 87 Proc. Natl. Acad. Sci. USA. 1889 (1990). After resuspension, the DNA was overlaid with one drop of mineral  
25 oil (Perkin-Elmer Cetus), denatured at 100°C for 5 minutes, and 0.5  $\mu\text{l}$  of 5 M NaCl was added. The DNAs were hybridized at 68°C for at least 20 hours.

To amplify the difference products (DPI), the 3'-recessed ends were first filled in with DNA polymerase. The hybridized product was diluted with 198  $\mu\text{l}$  of TE buffer, and a 50  $\mu\text{l}$  of filling-in reaction was set up: 30  $\mu\text{l}$  of diluted DNA hybrids, 10  $\mu\text{l}$  of 5X PCR  
30 buffer, 2  $\mu\text{l}$  of dNTP (5 mM for each dATP, dCTP, dGTP, and dTTP), and 8  $\mu\text{l}$  of  $\text{ddH}_2\text{O}$ . The reaction was equilibrated at 72°C, 3U of Taq DNA polymerase (Perkin-Elmer Cetus) and 0.75U of pwo DNA polymerase (Boehringer Mannheim) were added at 72°C for 10

minutes. The filled-in DNA hybrids were purified by a PCR purification spin column (Boehringer Mannheim).

To reduce the background for DPl amplification, mung bean nuclease treatment of the filled-in DNA hybrids was performed: 9  $\mu$ l of filled-in DNA hybrid solution mixed  
5 with 1  $\mu$ l of 10X mung bean nuclease reaction buffer and 50U of mung bean nuclease (Boehringer Mannheim) was incubated at 30°C for 30 minutes. The reaction was terminated by heat inactivation at 100°C for 5 minutes after adding 40  $\mu$ l of 50 mM Tris-HCl (pH 8.9).

To amplify DPl, 1 tube of 100  $\mu$ l PCR reaction was set up. The reaction  
10 consisted of: 20  $\mu$ l of mung bean nuclease-treated DNA hybrid solution, 20  $\mu$ l of 5XPCR buffer, 5  $\mu$ l of dNTP (5mM for each dATP, dCTP, dGTP, and dTTP), 2  $\mu$ l of the SEQ ID NO:3 adaptor, and 52  $\mu$ l of ddH<sub>2</sub>O. The reaction was hot started at 94°C for 2 minutes after adding 3U of Taq DNA polymerase (Perkin-Elmer Cetus) and 0.75U of pwo DNA polymerase (Boehringer Mannheim). The conditions for the PCR were 1 minute  
15 at 95°C, 1 minute at 67°C, and 3 minutes at 72°C for 35-40 cycles. The last cycle was followed with a 10 minutes extension step at 72°C. The DPl was purified by PCR purification spin column (Boehringer Mannheim), and the concentration of the DNA was determined at O.D.<sub>260</sub> in a spectrophotometer.

To prepare the second round of hybridization/subtraction and amplification steps,  
20 3  $\mu$ g of DPl was digested with the restriction endonuclease Msp I (10U/ $\mu$ g) in a 30  $\mu$ l reaction. The digested DP I was purified with PCR purification spin column (Boehringer Mannheim).

To put a new set of adapters (MSC24 having SEQ ID NO:5 and MSC12 having SEQ ID NO:6) on DPl, 2  $\mu$ l (0.1  $\mu$ g) of DPl was mixed with 1  $\mu$ l of the SEQ ID NO:5  
25 adaptor (0.6  $\mu$ g/ $\mu$ l), 1  $\mu$ l of the SEQ ID NO:6 adaptor (0.3  $\mu$ g/ $\mu$ l), 1  $\mu$ l of 10X ligation buffer (Boehringer Mannheim), and 6  $\mu$ l of ddH<sub>2</sub>O. The annealing and ligation reactions were done as before. The ligate was diluted with 40  $\mu$ l of TET buffer to the concentration of 2 ng/ $\mu$ l.

To set up the hybridization reaction, 20 ng (10  $\mu$ l) of diluted ligate was mixed with  
30 20  $\mu$ g (40  $\mu$ l) of driver amplicon DNA. The DNA mix was extracted with an equal volume of phenol/ chloroform once, and precipitated with 2.5 volumes of 100% ethanol after adding 1/4 volume of 10 M of NH<sub>4</sub>Ac in a dry ice/ethanol bath for 10 minutes. The

reaction solution was equilibrated to room temperature and the DNA was collected by centrifugation. The pellet was washed twice with 70% ice cold ethanol, dried, and resuspended in 2  $\mu$ l of 5XEE. 0.5  $\mu$ l of 5 M NaCl was added after the DNA mix was denatured at 100°C for 5 minutes. The hybridization reaction was carried out at 68°C for a minimum period of 20 hours.

To amplify the secondary different product (DP2), the hybrids were diluted with 198  $\mu$ l of TE buffer, and 100  $\mu$ l of PCR amplification reaction was set up as follows: 4  $\mu$ l of the diluted hybrid solution, 5  $\mu$ l of dNTP (5 mM for each dATP, dCTP, dGTP, dTTP), 20  $\mu$ l of 5X PCR reaction buffer, and 69  $\mu$ l of ddH<sub>2</sub>O. While the PCR tube was equilibrated at 72°C, 3U of Taq DNA polymerase (Perkin-Elmer Cetus) and 0.75U of pwo DNA polymerase (Boehringer Mannheim) were added to fill-in the 3'-recessed ends for 10 minutes. 2  $\mu$ l of the SEQ ID NO:5 adaptor (0.6  $\mu$ g/ $\mu$ l) was then added. The PCR reaction was not started by raising the temperature to 94°C for 2 minutes, and 35 to 40 cycles of PCR reaction were performed under the same conditions as were used for amplifying DP1.

DP2 usually contained several individual DNA fragments when electrophoresed on a 2 % agarose gel. The individual DNA fragments were purified by DNA gel extraction kit (Qiagen Inc.), and subcloned into pUC118 vector, which was linearized by the restriction endonuclease AccI. If the DP2 product had some DNA smear, it was treated with mung bean nuclease as described above.

For each MDD, twelve to twenty colonies were chosen to amplify their inserts, from which different sized probes were selected for amplicon Southern Blotting, and human genomic DNA Southern Blotting.

**Amplicon DNA Southern Blot.** The first round of positive probe screening was performed with amplicon DNA Southern Blots. Non-Radiation Southern Blot and Detection Kits (Genius™) were purchased from Boehringer Mannheim. Probe labeling, and detection followed the instruction of the manufacturer.

Two to three  $\mu$ g of tester and driver amplicon DNA were electrophoresed on a 2% agarose gel, and blotted to positively charged nylon membranes (Boehringer Mannheim). For prehybridization, the membranes were placed at 68°C for 2-4 hours in solutions containing 6XSSC, 5X Denhardt's solution, 0.5% SDS, 0.1 M EDTA, and 50

5  $\mu\text{g/ml}$  of salmon sperm DNA. Under the same conditions, the probes were added, and hybridized to the membranes overnight. The membranes were then rinsed three times with 2XSSC, 1X blot wash (12 mM  $\text{Na}_2\text{HPO}_4$ , 8 mM  $\text{NaH}_2\text{PO}_4$ , 1.4 mM  $\text{Na}_4\text{P}_2\text{O}_7$ , 0.5% SDS) at 68°C, and further washed three times (30 minutes for each time) with the same  
10 buffer at 68°C. Next, the membranes were equilibrated in buffer A (100 mM Tris.HCl, 150 mM, pH 7.5) and transferred into buffer B (2% black reagent in buffer A) solution and incubated at room temperature for one hour. The membranes were then washed 2 times (15 minutes for each time) with buffer A, and equilibrated in buffer C (100 mM Tris.HCl, 100 mM NaCl, 10 mM  $\text{MgCl}_2$ ). Before the membranes were exposed on  
15 Kodak X-OMAT film, they were rinsed in lumi-P530 for 1 min and kept in a plastic sheet protector. Positive clones were identified by observing the probe hybridized more heavily to test amplicon DNA than to driver amplicon DNA.

15 Human Genomic DNA Southern Blot. The positive probes that were confirmed by the amplicon DNA Southern Blot experiment were tested further by human genomic DNA Southern Blotting. Genomic DNAs isolated from cancer tissues, and their respective matched normal tissues, were digested with Msp I restriction endonuclease, electrophoresed on 1.5% agarose gels, and transferred to Hybond Membranes (Amersham, Arlington Heights, IL). These membranes were exposed to UV light to  
20 immobilize the DNA. The probes for the Southern Blot were labeled with High Prime DNA labeling kits (Boehringer Mannheim) following the instructions of the manufacturer. The procedure for hybridization and blot wash were the same as in the Amplicon DNA Southern Blotting section.

### EXAMPLE 2

#### Different methylation patterns at CpCpG sites in malignant B-cells and neutrophil cells

DNA isolated from malignant B-cells and neutrophil cells of four CLL patients  
5 (#58, #111, #112, #128) was used to perform MDD as described above in the Materials  
and Methods provided in Example 1. The DNA from leukemic B-cells was used as the  
tester, and the DNA from neutrophil cells was used as the driver. Tester and driver DNAs  
were simultaneously digested with the restriction endonucleases, Msp I and Mse I. After  
ligating the first set of adapters to the digested tester and driver DNA fragments, tester  
10 and driver amplicons were generated by PCR amplification (see Materials and Methods).  
After two rounds of hybridization/ subtraction and amplification, the different products  
(DP2) which were isolated from patient 111 appeared as individual DNA fragment bands  
on an agarose gel (Fig. 2). The DP2 fragments were then subcloned into the plasmid  
pUC118 (see Materials and Methods), and amplified by PCR. For the first round of  
15 selection, inserts of different sizes were selected and hybridized to the tester and driver  
amplicon DNAs from which they were derived in a non-radioactive Southern Blot  
experiment (see materials and methods). The positive probes were selected if they  
hybridized a single band in the tester amplicon only, or if they hybridized to both tester  
and driver amplicon DNA, but the hybridization to the driver amplicon DNA band was  
20 far less intense. The final proof of a positive probe was that it displayed a different  
methylation pattern in a genomic DNA Southern Blot experiment in which the parental  
malignant B-cell, and normal neutrophil cell genomic DNA, were digested with the  
restriction endonuclease Msp I (see materials and methods).

The CLL58 probe isolated from patient #111 fulfilled these requirements. When  
25 CLL58 was hybridized to its tester and driver amplicon DNA, the tester amplicon DNA  
gave a heavy hybridization band, while the driver amplicon DNA gave a much less intense  
hybridized band (Fig.3).

The CLL58 probe was further tested by hybridization to a genomic Southern blot  
of MspI-digested malignant B-cell and neutrophil cell DNA. The genomic Southern Blot  
30 revealed that a larger fragment was much less intensely hybridized than a lower molecular  
weight fragment in the malignant B-cell DNA, while in the neutrophil cell DNA, just the  
opposite result occurred (Fig. 4). These results demonstrate that in human genomic DNA,

there are methylated external cytosine residues at CpCpG sequences, and those CpCpG sequences are differentially methylated in different types of cells. Moreover, the CLL58 probe can detect and distinguish cell-type specific methylation patterns.



### EXAMPLE 3

#### MDD-Isolated Probes from Human Breast Cancer Tumors Also Detect Ovarian and Colon Cancers

We have proven the existence of methylation of the external cytosine residue at CpCpG sites in the human genome, and have isolated CpCpG site related methyl-polymorphic markers in a working model system. MDD was then applied to successfully isolate CpCpG related methyl-polymorphic markers from human breast cancer biopsies.

#### Isolation of Probes From Breast Tumor DNA:

MDD was used to test five pairs of DNA samples (tumor DNA and matched normal DNA) obtained from five breast cancer patients. To isolate methyl-polymorphic markers, tumor DNA served as the tester, and normal DNA served as the driver. MDD was performed in the same manner as described in Example 1. Individual DNA fragments were isolated by MDD from three breast cancer patients after two rounds of hybridization/subtraction and amplification (Fig. 5).

Three probes were further analyzed, BR50, BR104, and BR254, isolated from patient #14, patient #13, and patient #4, respectively. Each of these three probes appropriately hybridized with their own tester amplicon DNA, but also hybridized slightly with the driver amplicon DNA (Fig. 6a-c).

These probes were then examined further by hybridizing them to *MspI*-digested parental tumor and matched normal genomic DNAs. The genomic Southern Blot for probe BR50 showed that the probe hybridized with a lower band to a much greater degree than to an upper band in tumor DNA (Fig. 7a). However, in normal DNA, just the opposite occurred -- the BR50 probe hybridized to the upper band more than to the lower band (Fig. 7a).

The genomic DNA Southern Blot for probe BR104 showed that the probe hybridized with a lower and a middle band in the tumor DNA, while in the normal DNA, probe BR104 hybridized with a middle and an upper band (Fig. 7b).

For the probe BR254, the Southern Blot result showed that it hybridized with a lower band and a upper band in the tumor DNA, however, it only hybridized with an upper band in the normal DNA (Fig. 7c).

The results of three cases, indicate that in normal cells the genomic DNA is more highly methylated, is cut less frequently and gives rise to larger fragments than is observed

for DNA from tumor cells. The isolated probes therefore hybridize to higher molecular weight bands in normal DNA. In contrast, the lower bands which were hybridized with probes BR50, BR104, and BR254, are generated by demethylation of the external C residue at CpCpG DNA sites in the tumor cells. The weakly hybridized upper bands in the tumor DNA are likely caused by normal cell contamination

#### **Breast Tumor Probes Detect Hypomethylation in Other Cancer Patients**

To understand whether this hypomethylation event also happens in other breast cancer patients, normal cell and tumor cell DNA was isolated from ten breast cancer patients, cleaved by the Msp I enzyme, and probed with BR50 and BR104. Of the ten breast patients whose DNA was probed with BR50, five had a hybridization pattern which was similar to that of the patient from which BR50 was isolated (the "parental patient") (Fig. 8). Of the ten breast patients whose DNA was probed with BR104, five had a hybridization pattern similar to that of the parental patient. (See Fig. 9, providing four of these five patterns).

The BR254 probe was hybridized to normal and tumor DNA of 16 breast cancer patients. Of the 16 patients, 10 had a hybridization pattern which was similar to that of the parental patient. However, some differences in the methylation patterns of these patients were observed. Some patients only had the lower band in their tumor DNA, and the upper band in their normal DNA. (Fig. 10) However, other patients had the lower band in their tumor DNA, but had both lower and upper bands in their normal DNA (Fig. 10). These results indicate that the external cytosine residue at CpCpG sites are fully demethylated in some breast tumors. However, the lower bands appearing in normal DNAs probably were caused by a partial demethylation event, which suggests that the normal cells were under the process of neoplastic change. In general, the different extent of demethylation identified by probes BR50, BR104, and BR254 in human breast cancer patients is a common phenomena and may be used to predict the extent of neoplastic change.

**Detection of a DNA Amplification Event in a Breast Cancer Patient** The BR254 probe detected a highly amplified DNA event in one of the 16 breast cancer patients tested with that probe (Fig. 11). This result provides evidence that imbalanced DNA methylation

may cause a conformational change of chromatin which then induces a DNA amplification.

**Hypomethylation Detection by BR50, BR104, and BR254 in Ovarian Cancers.** To  
5 examine whether the same hypomethylation phenomena also happens in human ovarian  
cancer, DNA samples isolated from 8 patients with ovarian cancer were analyzed. For  
some patients, two DNA samples were obtained, one from primary ovarian cancer tissue,  
and the other from their matched normal tissue. For other patients, three DNA samples  
were obtained, they were from primary ovarian cancer tissue, metastatic tumor tissue, and  
10 matched normal tissue. The DNAs were cleaved with *MspI*, and hybridized with the  
probes BR50 and BR104 in Southern Blot experiments.

Of the eight ovarian patients tested with the BR50 probe, five patients had similar  
hybridization patterns: the BR50 probe hybridized almost equally to lower and upper  
bands in their normal DNA samples, while in the primary tumor DNA, the upper band was  
15 only slightly hybridized. However, in the metastatic tumor DNA of these 5 patients, only  
the lower band was hybridized (Fig. 12).

No different hybridization patterns were detected using the probe BR104. This  
result indicates that the BR104 probe was detecting a tissue-specific hypomethylation  
event.

20 Eleven patients' DNA samples were examined using the BR254 probe. The  
Southern Blot results showed that of 11 patients, 4 had the same hybridization patterns  
which appeared in the parental DNA pair (Fig. 13)

These results suggest that probes BR50 and BR254 detect tumor-related  
hypomethylation events in both breast and ovarian cancer cells. Interestingly, in the case  
25 of probe BR50, the complete disappearance of the upper band in the metastatic tumor  
DNA indicates that the extent of demethylation seems to be related to the progression of  
the disease.

**Hypomethylation Detection using BR50, BR104, and BR254 on Colon Cancer DNA.**

5 The methylation patterns of tumor and matched normal DNA samples isolated from 10 colon cancer patients were examined using BR50, BR104, and BR254 probes. DNAs were digested by the MspI enzyme, Southern Blots prepared and the blots were hybridized with these probes.

No difference in the methylation patterns of colon tumor and normal DNAs were detected by probes BR50 and BR104 (data not shown), indicating that both probes were detecting a tissue-specific hypomethylation event in colon cells.

10 However, for the probe BR254, different methylation patterns were observed. Of the ten colon cancer patients tested, six had hypomethylation patterns opposite to those observed for the breast and ovarian cancer patients. Instead of hybridizing more to an upper band in the normal DNA samples, probe BR254 hybridized with the lower band. Moreover, instead of hybridizing more to the lower band in the tumor DNAs, BR254  
15 hybridized equally with the lower band and a upper band (Fig. 14). These results demonstrate that while the BR254 probe detects hypomethylation in breast cancer cells, in colon cancer DNA the BR254 probe detects hypermethylation. Therefore, methylation patterns are not only tissue-type specific, but cancer-specific, and can be different in different types of cancers.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Duffy, Hao-peng X

(ii) TITLE OF INVENTION: Detection of Differentially Methylated Nucleic Acids

(iii) NUMBER OF SEQUENCES: 10

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Kenyon &amp; Kenyon

(B) STREET: One Broadway

(C) CITY: New York

(D) STATE: New York

(E) COUNTRY: U.S.A

(F) ZIP: 10004

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Tsevdos, Estelle J.

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212)425-7200

(B) TELEFAX: (212)425-5288

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCGTCGTCA GGTCAGTGCT TCAC

24

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGTGAAGCA CT

12

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TAGAGCCACG TAGCTGCTGT AGTC

24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGACTACAG CA

12

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACCGTGGACT GGATAGGTTC AGAC

24

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 12 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGTCTGAAC CT

12

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 309 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

## (iii) MOLECULE TYPE: DNA (cDNA)

## (iv) SEQUENCE DESCRIPTION: SEQ ID NO: 7

```

CCGGGGCCCC ACAGGCCCCG TGCTGAAACA GGGCTGGAAC CAGGCAACTT GTTCTTTTGT   60
CCTGACATCC CTGCGCAGCT GGTACCGAGT GGACAGCCCC AACAACTCA TGAACCCGCT   120
GGTCGCTGGG GTCTTCGGAG CCATTGTGGG AGCGGCCAGT GTCTTCGGAA ATGCTCCTCT   180
GCACGAGATC GAGCCCCAGA TCGGGGACCT GGAGGTGCAC AAATGCATAA CACATGGGAC   240
TGTGGCTGCA AATCCTGAGG GAAGGGCACA AGACCTTCCT ACAAGGGCAC TATTGCCCCG   300
CTGGGCCCGG                                     309

```

## (2) INFORMATION FOR SEQ ID NO: 8.

## (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 332 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (cDNA)

## (iv) SEQUENCE DESCRIPTION: SEQ ID NO: 8.

```

CCGGGCAATC TCCACAGGCA CTCCTCCAC TGTCTGTACA ACGTCCTTAG CTCTTCCTTC   60
TGCCAGTGT GCAATCTTCC CAGATGGGCC TTCTGCCACA CATTACGTC TGATCTGGCA   120
GGTGTTCCTT CTGCCAAACC TTCATCCTTC TCGGATGCTC TTCTGCCACA TGTGCAATCC   180
TGTCAGACGT TCCTGCCACT CGAGGTAGCT GGTCTGATGT GAGCATGGAA CCAGGGGGTC   240
CCCCTACCAC CACCAGGAAT GTCAGATGAT CATGAGGTGT CGGTTGGGCG GTTCTTGAAC   300
CCTATCTATG GAGGATAACC GCTCACGGCC GG                                     332

```

## (2) INFORMATION FOR SEQ ID NO: 9.

## (i) SEQUENCE CHARACTERISTICS



- (A) LENGTH: 568 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO 9

```

CCGGCCTGGG AGGAGGGTTC CCACCAGGCT TTCTCAAGCC CATCACAAGT GAGTCAGGAA   60
CCATCTCCCA TTGGTCCTGG TAAGGAGGAG ATTTGGAGGA AGGACTGGGA AGCCTGTGGA   120
CGGGGAGGGT ACTGGGGGTG GAAGTGGAAG AAGGTGGCAC AGTAACAGAC TCCCTTCTCC   180
TGAGATTCCC AGCATGGATT GGAGGGGCTG CCCTGCAGCG TCCCTTACCC CTTGTTACCT   240
GGCAGCCTGC AAGTAGCTCT AGGGCCAGCC ATGTGGTCAC CTACATCTGT TGGGGGAGAG   300
AAGGAAAACA GATGCCCCGA GTCGTAGAGA CTTGCATTAT CAGCCCCTGG GTCCAACAGC   360
AAGTGAGCTA AGCACTTGCA AATGTTATCT GATGTATTCT TTATAAGAAG GGGTCACAAA   420
ACTGAGACTT AGAACATGTA AGAGGGAGAC AGGGATTGTA GCCTGGATCA GCTATGTGAG   480
GATCTECTGT CTGTCCCCGC TGGTTCCAAT TCAGGTCTTA GAGTAAAAGG GTTGGGGATG   540
AGCTCTCTGG CAGGACAGTG CCCTCCGG

```

568

(2) INFORMATION FOR SEQ ID NO 10

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 1005 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

CCGGAGGCAG GCACAGGACT CGGGAGGGAC GCTGCCAGCT CTCTGGGTGC TGAGTTCACA   60

```

AGGCTGCATT CATGATTTTC AATAGACCTG TGATGGTCTG TGCCCAGTGC TGGGGACACA 120  
 GAAGAGTCAA ACCTGGCTCC TGACCTGGAC CTGGATCATC ACGTGACAGG GAGGAGAGCG 180  
 ATCCAGGCTG ATGAGGAAAG CGCATGACAT GGGGTCTTAG GAGCAGTGAG GGGCAGAGCC 240  
 ATGGCCAAAG GCCCCGCCAT GGAAGCTGAG GACTCTGGCA CCAGATGGAG GCAGTTGACC 300  
 GACCTCTGCC CTTGGGGTCC AACCCATGGG CTTCTCATAC ATAGGGGTGA AAAAGGCCAT 360  
 TCTATTTATG CAGAATTTTC CCATGTGGCC AGGCAGCAGA AGTCCAGAGG GGTAGGGGCC 420  
 ACTCAGGGTC ACACAGAACA GCAGTTGCTG AAGACTGGGG AAGTCCAGGC CTAGGCTCCA 480  
 CCTGCCCTTC CCCTGACATG GGGCCACCAC TAGCCTTTTA TGGGCAGGCC TGGCTGCTGG 540  
 TGGTTGGAAT AACATCTGAC TCCAGTGGGT GTCTGTCACC GTCTCCAGAC AGGAGACAGA 600  
 GACAGAGGGT CAAAGTTCAC TATGGCTCTT TGGGGCAATG AAATGCTGTG TTCTAGCCTC 660  
 TTGCCAGAAA TCAGCCAAAG TCAAGGAAAG CCTGACTCCC ACAGTTATCA CAGAAAGAGC 720  
 ACCCACTTTC CAGCCCAGAC AGCTGCACCC CAGCTGGGTC CTGGCAGCCC CAGCTTCAGC 780  
 CTGGGCGGT A TGTTCAGGC CCTCGATCA TCTGACCCTA ATATCACCCC TTCACACCCC 840  
 CTCCACTTTC TCGGGGAGCC ACCCCGAACC TTTGAATGGG GGAGATCCTG GAGGCTCTGC 900  
 AATTTTCAGT GTAAACTGCC TGGAGTTCCC CACTTCACCC TCATCTGGTT CACCTGTGGA 960  
 CTCCCAACAG AGCAGGCCCA GGAAACGCGG GGCCTCTGAG GCCGG 1005

WHAT IS CLAIMED:

1. A method of detecting whether a CNGtriplet is hypomethylated or hypermethylated in a genomic DNA present in a test sample of cells which comprises:
  - a) cleaving genomic DNA isolated from a control sample of cells and a test sample  
5 of cells with a master restriction enzyme to generate a cleaved control-cell DNA and a  
cleaved test-cell DNA;
  - b) hybridizing a probe to said cleaved control-cell DNA and said cleaved test-cell  
DNA to form a control-hybridization complex and a test-hybridization complex, wherein  
said probe comprises a nucleic acid fragment which is methylated in said control sample  
10 of cells but which is not methylated in said test sample of cells; and
  - c) determining whether the size of the control-hybridization complex is the same  
as the size of the test-hybridization complex;  
wherein said master restriction enzyme cleaves a nonmethylated CNGDNA  
sequence but does not cleave a methylated CNGDNA sequence.
- 15 2. The method of Claim 1 wherein said nucleic acid fragment is cut by said  
master restriction enzyme in genomic DNA from said control sample of cells but which  
not cut by said master restriction enzyme in genomic DNA said test sample of cells.
3. The method of Claim 1 wherein said probe comprises a DNA selected from  
the group consisting of SEQ ID NO:7-10.
- 20 4. The method of Claim 1 wherein said master restriction enzyme is MspI,  
BsiSI or HinfI.
5. The method of Claim 1 wherein said test sample of cells comprises  
cancerous cells.
6. The method of Claim 5 wherein said cancerous cells are isolated from a  
25 person with Wilms' cancer, breast cancer, ovarian cancer, colon cancer, kidney cell  
cancer, liver cell cancer, lung cancer, leukemia, rhabdomyosarcoma, sarcoma, or  
hepatoblastoma.
7. The method of Claim 1 wherein said control sample of cells comprises non-  
cancerous cells derived from the same person as said test sample of cells.
- 30 8. The method of Claim 1 wherein said control sample of cells comprises non-  
cancerous cells derived from the same tissue-type as said test sample of cells.

9. A method of isolating a probe to detect hypomethylation in a CNGtriplet of DNA which comprises:

a) cleaving a tester sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

5 b) amplifying said cleaved tester sample using primers that recognize DNA ends cut by said master restriction enzyme, to generate a first tester amplicon;

c) melting and hybridizing said first test amplicon with a driver DNA to generate a tester-tester product and a tester-driver product;

10 d) adding nucleotides to the ends of said tester-tester product and said tester-driver product to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

e) amplifying said blunt-ended tester-tester product and said blunt-ended tester-driver product by using primers which recognize only master-enzyme-cut DNA ends, to produce a second tester amplicon;

15 f) isolating a discrete DNA fragment from said second tester amplicon as a probe to detect said hypomethylation or said hypermethylation in a CNGtriplet of DNA;

wherein said master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence;

20 wherein said partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of said genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein said driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by said master restriction enzyme.

25 10. A method of isolating a probe to detect hypomethylation in a CNGtriplet of DNA which comprises:

a) cleaving a tester sample of genomic DNA with both a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

30 b) ligating a first set of adaptors onto master enzyme cut DNA ends of said cleaved tester sample to generate a first-tester amplification template;

c) amplifying said first-tester amplification template to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to said first set of adaptors;

5 d) cleaving off said first adaptors from said first-tester amplicon and ligating a second set of adaptors onto DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends;

e) melting and hybridizing the second-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a mixture of tester-tester product and tester-driver product;

10 f) adding nucleotides onto DNA ends present in said mixture to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

g) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-tester amplicon;

15 f) isolating a discrete DNA fragment from said second tester amplicon as a probe to detect said hypomethylation or said hypermethylation in a CNGtriplet of DNA;

wherein said master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence;

20 wherein said partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of said genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein said driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme and amplified using primers that recognize DNA ends cut by said master restriction enzyme.

25 11. The method of Claim 9 or 10 which further comprises digesting said blunt-ended tester-tester product and said blunt-ended tester-driver product with a nuclease to remove unhybridized regions in the blunt-ended tester-tester product and the blunt-ended tester-driver product.

30 12. The method of Claim 10 which further comprises cleaving off said second adaptors from said second-tester amplicon and ligating a third set of adaptors onto DNA ends of the second-tester amplicon to generate a third-adaptor-tester which has third-adaptor ends;

melting and hybridizing the third-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a second mixture of second tester-tester product and second tester-driver product,

5 adding nucleotides onto DNA ends present in said second mixture to make a second blunt-ended tester-tester product and a second blunt-ended tester-driver product;

amplifying the second blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to third adaptor ends to generate a third-tester amplicon; and

10 isolating a discrete DNA fragment from said third tester amplicon as a probe to detect said hypomethylation in a CNGtriplet of DNA.

13. The method of Claim 9 or 10 wherein said master restriction enzyme is MspI, BsiSI or Hin2I.

14. The method of Claim 9 or 10 wherein said partner restriction enzyme is MseI, Sau3A, RsaI, TspEI, MaeI, NlaIII or DpnI.

15 15. The method of Claim 10 wherein said first set of adaptors cannot be ligated onto DNA ends cut by said partner enzyme.

16. The method of Claim 10 wherein said second set of adaptors cannot be ligated onto DNA ends cut by said partner enzyme.

17. The method of Claim 10 wherein said first set of adaptor oligonucleotides are selected from the group consisting of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:6

18. The method of Claim 10 wherein said second set of adaptor oligonucleotides are selected from the group consisting of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:6.

25 19. The method of Claim 9 wherein said first tester amplicon is melted and hybridized with about a 10-fold to about a 10,000-fold molar excess of said driver DNA.

20. The method of Claim 9 wherein said first tester amplicon is melted and hybridized with about a 10 to about a 1000-fold molar excess of a driver DNA.

30 21. The method of Claim 9 wherein said first tester amplicon is melted and hybridized with about a 100-fold molar excess of a driver DNA.

22. The method of Claim 10 wherein said second-adaptor-tester is melted and hybridized with about a 100-fold molar excess of a driver DNA.

23. The method of Claim 9 or 10 wherein isolating said discrete DNA fragment comprises cloning said discrete DNA fragment into a vector.

24. The method of any one of Claims 9 or 10 wherein said tester sample of genomic DNA is isolated from a sample of cancerous tissue.

5 25. The method of Claim 24 wherein said cancerous tissue is isolated from a patient with Wilms' cancer, breast cancer, ovarian cancer, colon cancer, kidney cell cancer, liver cell cancer, lung cancer, leukemia, rhabdomyosarcoma, sarcoma, or hepatoblastoma.

26. The method of Claim 9 or 10 wherein said tester sample of genomic DNA is isolated from a patient with a genetic defect in genomic imprinting.

10 27. The method of Claim 9 or 10 wherein said driver DNA is genomic DNA isolated from normal cells of the same person as said tester DNA.

28. The method of Claim 9 or 10 wherein said driver DNA is genomic DNA isolated from normal cells of the same tissue type as said tester sample of genomic DNA.

15 29. A method of detecting whether a CNGtriplet of DNA is hypomethylated in a tester sample of genomic DNA which comprises:

a) cleaving a tester sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

20 b) amplifying said cleaved tester sample by in vitro DNA amplification using primers that recognize DNA ends cut by said master restriction enzyme, to generate a first tester amplicon;

c) melting and hybridizing said first test amplicon with a driver DNA to generate a tester-tester product and a tester-driver product.

25 d) adding nucleotides to the ends of said tester-tester product and said tester-driver product to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;

e) amplifying said blunt-ended tester-tester product and said blunt-ended tester-driver product by in vitro DNA amplification using primers which recognize only master-enzyme-cut DNA ends, to produce a second tester amplicon;

30 f) isolating a discrete DNA fragment from said second tester amplicon;

j) preparing a probe from said discrete DNA fragment;

g) isolating genomic DNA from a control tissue sample and a test tissue sample to generate a control-tissue DNA and a test-tissue DNA;

h) cleaving said control-tissue DNA and said test-tissue DNA with said master restriction enzyme to generate cleaved control-tissue DNA and cleaved test-tissue DNA;

i) hybridizing said probe to said cleaved control-tissue DNA and said cleaved test-tissue DNA to form a control-hybridization complex and a test-hybridization complex; and

5 j) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex and thereby detecting hypomethylation in a CNGtriplet of DNA which is present in a tester sample of genomic DNA;

wherein said master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence;

10 wherein said partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of said genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein said driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme and amplified using primers that recognize DNA ends cut by  
15 said master restriction enzyme.

30. A method of detecting hypomethylation or hypermethylation in a CNGtriplet of DNA which is present in a test tissue sample which comprises:

a) cleaving a tester sample of genomic DNA with both a master restriction enzyme and a partner restriction enzyme to generate a cleaved tester sample;

20 b) ligating a first set of adaptors onto DNA ends of said cleaved tester sample to generate a first-tester amplification template;

c) amplifying said first-tester amplification template to generate a first-tester amplicon by in vitro DNA amplification using primers that hybridize to said first set of adaptors;

25 d) cleaving off said first adaptors from said first-tester amplicon and ligating a second set of adaptors onto DNA ends of the first-tester amplicon to generate a second-adaptor-tester which has second adaptor ends;

e) hybridizing the second-adaptor-tester with about a 10-fold to about a 10,000-fold molar excess of a driver DNA to generate a mixture of tester-tester product and  
30 tester-driver product;

f) adding nucleotides onto DNA ends present in said mixture to make a blunt-ended tester-tester product and a blunt-ended tester-driver product;



g) amplifying the blunt-ended tester-tester product and the blunt-ended tester-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-tester amplicon;

h) isolating a discrete DNA fragment from said second-tester amplicon;

5 i) preparing a probe from said discrete DNA fragment;

j) isolating genomic DNA from a control tissue sample and a test tissue sample to generate a control-tissue DNA and a test-tissue DNA;

k) cleaving said control-tissue DNA and said test-tissue DNA with said master restriction enzyme to generate cleaved control-tissue DNA and cleaved test-tissue DNA;

10 l) hybridizing said probe to said cleaved control-tissue DNA and said cleaved test-tissue DNA to form a control-hybridization complex and a test-hybridization complex; and

n) observing whether the size of the control-hybridization complex is the same as the size of the test-hybridization complex;

15 wherein said master restriction enzyme cleaves a nonmethylated CNGDNA sequence but does not cleave a methylated CNGDNA sequence;

wherein said partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of said genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

20 wherein said driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme and amplified using primers that recognize DNA ends cut by said master restriction enzyme.

31. The method of Claim 29 or 30 wherein said method is used for detecting cancer, cell-type specific expression, tissue-specific expression, or a genetic disease which alters methylation.

25 32. The method of Claim 31 wherein said cancer is Wilms' cancer, breast cancer, ovarian cancer, colon cancer, kidney cell cancer, liver cell cancer, lung cancer, leukemia, rhabdomyosarcoma, sarcoma, or hepatoblastoma.

33 The method of Claim 31 wherein said genetic disease which alters methylation is Beckwith-Wiedemann syndrome or Prader-Willi syndrome.

30 34. The method of Claim 29 or 30 wherein said master restriction enzyme is MspI, BsiSI or Hin2I.

35. The method of Claim 29 or 30 wherein said partner restriction enzyme is MseI, Sau3A, RsaI, TspEI, MaeI, NlaIII or DpnI.

36. The method of Claim 30 wherein said first set of adaptors cannot be ligated onto DNA ends cut by said partner enzyme.

5 37. The method of Claim 30 wherein said second set of adaptors cannot be ligated onto DNA ends cut by said partner enzyme.

38. The method of Claim 30 wherein said first set of adaptors are selected from the group consisting of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:6.

10 39. The method of Claim 30 wherein said second set of adaptors is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:6.

40. The method of Claim 29 wherein said first test amplicon is melted and hybridized with about a 10-fold to about a 10,000-fold molar excess of said driver DNA.

15 41. The method of Claim 29 wherein said first test amplicon is melted and hybridized with about a 100-fold molar excess of said driver DNA.

42. The method of Claim 29 wherein said blunt-ended tester-tester is hybridized with about a 10-fold to about a 10,000-fold molecular excess of said driver DNA.

20 43. The method of Claim 29 wherein said blunt-ended tester-tester is hybridized with about a 100-fold molecular excess of said driver DNA.

44. The method of Claim 29 or 30 wherein isolating said discrete DNA fragment comprises cloning said discrete DNA fragment into a vector.

25 45. The method of Claim 29 or 30 wherein said driver DNA is genomic DNA isolated from normal cells of the same person as said tester DNA.

46. The method of Claim 29 or 30 wherein said driver DNA is genomic DNA isolated from normal cells of the same tissue type as said tester sample of genomic DNA.

30 47. The method of Claim 29 or 30 wherein observing whether the size of said control-hybridization complex is the same as the said size of the test-hybridization complex is done by Southern blot analysis.

48. The method of Claim 29 or 30 which further comprises digesting said blunt-ended tester-tester product and said blunt-ended tester-driver product with a

nuclease to remove unhybridized regions in the blunt-ended tester-tester product and the blunt-ended tester-driver product.

49. A method of isolating a probe to detect hypermethylation in a CNGtriplet of DNA which comprises:

- 5 a) cleaving a normal sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved normal sample;
- b) amplifying said cleaved normal sample using primers that recognize DNA ends cut by said master restriction enzyme, to generate a first normal amplicon;
- c) melting and hybridizing said first normal amplicon with a test driver DNA to  
10 generate a normal-normal product and a normal-driver product;
- d) adding nucleotides to the ends of said normal-normal product and said normal-driver product to make a blunt-ended normal-normal product and a blunt-ended normal-driver product;
- e) amplifying said blunt-ended normal-normal product and said blunt-ended  
15 normal-driver product by using primers which recognize only master-enzyme-cut DNA ends, to produce a second normal amplicon;
- f) isolating a discrete DNA fragment from said second normal amplicon as a probe to detect said hypermethylation in a CNGtriplet of DNA;
- wherein said master restriction enzyme cleaves a nonmethylated CNGDNA  
20 sequence but does not cleave a methylated CNGDNA sequence;
- wherein said partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of said genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and
- wherein said test driver DNA is cut with both the master restriction enzyme and  
25 the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by said master restriction enzyme.

50. A method of isolating a probe to detect hypermethylation in a CNGtriplet of DNA which comprises:

- 30 a) cleaving a normal sample of genomic DNA with a master restriction enzyme and a partner restriction enzyme to generate a cleaved normal sample;
- b) ligating a first set of adaptors onto master enzyme cut DNA ends of said cleaved normal sample to generate a first-normal amplification template;

c) amplifying said first-normal amplification template to generate a first-normal amplicon by in vitro DNA amplification using primers that hybridize to said first set of adaptors;

5 d) cleaving off said first adaptors from said first-normal amplicon and ligating a second set of adaptors onto DNA ends of the first-normal amplicon to generate a second-adaptor-normal DNA which has second adaptor ends;

e) melting and hybridizing the second-adaptor-normal DNA with about a 10-fold to about a 10,000-fold molar excess of a test driver DNA to generate a mixture of normal-normal product and normal-driver product;

10 f) adding nucleotides onto DNA ends present in said mixture to make a blunt-ended normal-normal product and a blunt-ended normal-driver product;

g) amplifying the blunt-ended normal-normal product and the blunt-ended normal-driver product by in vitro DNA amplification using primers that hybridize to second adaptor ends to generate a second-normal amplicon;

15 h) isolating a discrete DNA fragment from said second-normal amplicon as a probe to detect hypermethylation in a CNGtriplet of DNA;

wherein said master restriction enzyme cleaves a nonmethylated CCGG DNA sequence but does not cleave a methylated CNGDNA sequence;

20 wherein said partner restriction enzyme cleaves DNA to produce DNA fragments with a complexity of about 5% to about 25% of said genomic DNA in a size range which can be amplified by a DNA amplification enzyme; and

wherein said test driver DNA is cut with both the master restriction enzyme and the partner restriction enzyme, and then amplified by using primers that recognize DNA ends cut by said master restriction enzyme.

25 51. The method of Claim 49 or 50 wherein said master restriction enzyme is MspI, BsiSI or Hin2I.

52. The method of Claim 49 or 50 wherein said partner restriction enzyme is MseI, Sau3A, RsaI, TspEI, MaeI, NlaIII or DpnI.

30 53. The method of Claim 50 wherein said first set of adaptors cannot be ligated onto DNA ends cut by said partner enzyme.

54. The method of Claim 50 wherein said second set of adaptors cannot be ligated onto DNA ends cut by said partner enzyme.